

Санкт-Петербургский государственный университет

ШАРЫГИНА Софья Сергеевна

Выпускная квалификационная работа

**Клаузы с формами глагола быть в русской устной речи: корпусное
исследование**

Уровень образования: магистратура

Направление 45.04.02 «Лингвистика»

Основная образовательная программа ВМ.5624. «Теория и история языка и
языки народов Европы»

Профиль «Психо- и социолингвистика»

Научный руководитель:

доцент, Кафедра общего языкознания
имени Л.А. Вербицкой,
Риехакайнен Елена Игоревна

Рецензент:

Краснощёкова Софья
Викторовна, научный сотрудник,
Отдел теории
грамматики, Федеральное
государственное бюджетное учреждение
науки Институт лингвистических
исследований Российской академии наук

Санкт-Петербург

2021

Оглавление

Введение	4
Глава 1. Теоретические предпосылки исследования	6
1.1. О понятии «валентность»	6
1.2. Семантическая валентность, предикат, семантический актант	8
1.3. Синтаксическая валентность, актант, сирконстант, распространитель	12
1.4. Некоторые отклонения от прототипической ситуации	16
1.5. Семантическая роль, классификации семантических ролей	18
1.5.1. Понятие и общие принципы выделения семантических ролей	18
1.5.2. Различные инвентари семантических ролей	23
1.5.3. Определение семантической роли актантов: типичные проблемы и подходы к их решению	27
1.6. Специфика глагола <i>быть</i>	30
Глава 2. Анализ употребления клауз с глаголом <i>быть</i> в русской устной речи	32
2.1. Контекст и прикладные задачи исследования	32
2.2. Опорный перечень способов кодирования семантических ролей в русской устной речи	33
2.3. Общая информация о материале исследования	34
2.3.1. Корпус русской устной речи и работа с ним	34
2.3.2. Пользовательский подкорпус НКРЯ	37
2.3.3. Итоговый рабочий корпус	37
2.4. Модификации инвентаря семантических ролей	40
2.4.1. Субъект экзистенциального предложения	40
2.4.2. Объект обладания	42
2.4.3. Определение, Определение: как, Состояние, Роль и Функция	44
2.4.4. Тождественность	47
2.5. Частотность активации семантических ролей	47
2.5.1. Общая информация	47
2.5.2. Сочетаемость семантических ролей	49
2.6. Распределение семантических ролей по формам глагола <i>быть</i>	63

2.6.1. Типичные случаи активации семантических ролей в клаузах с «быть»	64
2.6.2. Типичные случаи активации семантических ролей в клаузах с «есть»	65
2.6.3. Типичные случаи активации семантических ролей в клаузах с формами будущего времени «быть»	66
2.6.4. Типичные случаи активации семантических ролей в клаузах с формами прошедшего времени «быть»	67
2.7. Способы кодирования наиболее абсолютно частотных семантических ролей в клаузах с глаголом <i>быть</i>	68
Заключение	72
Список использованной литературы	74
Приложение 1. Опорный инвентарь семантических ролей и способов их выражения	79
Приложение 2. Опорный инвентарь семантических ролей и способов их выражения: влияние конструкций	84
Приложение 3. Распределение по формам глагола <i>быть</i> : Экзистенциальный субъект	86
Приложение 4. Распределение по формам глагола <i>быть</i> : Место	87
Приложение 5. Распределение по формам глагола <i>быть</i> : Посессор	88
Приложение 6. Распределение по формам глагола <i>быть</i> : Объект Обладания	89
Приложение 7. Распределение по формам глагола <i>быть</i> : Определение	90
Приложение 8. Распределение: когда по формам глагола <i>быть</i> : Время: когда	91
Приложение 9. Распределение по формам глагола <i>быть</i> : Тождественность	92
Приложение 10. Распределение по формам глагола <i>быть</i> : <i>Агенс</i>	93

Введение

В работе представлено исследование, посвященное анализу клауз с глаголом *быть* в русской устной речи. **Целью** работы было описание семантических и синтаксических аспектов функционирования глагола *быть* посредством создания «инвентаря» семантических ролей, которыми могут обладать его актанты и сирконстанты, а также списка наиболее частотных способов кодирования этих ролей в русской устной речи.

Актуальность данной работы обусловлена, во-первых, сравнительной непредставленностью глагола *быть* в научной литературе при его чрезвычайно высокой частотности в письменной и устной речи и разнообразии грамматических функций, выполняемых им в русском языке. Во-вторых, настоящая работа является частью выполняемого исследователями из СПбГУ проекта «Алгоритмы восстановления редуцированных форм: роль системы языка», цель которого – разработать алгоритм взаимодействия различных уровней языка в процессе восприятия речи, который был бы близок к тому, как распознает речь носитель языка (проект поддержан грантом РФФИ №19-012-00629, рук. – Е. И. Риехайнен). Результаты исследования, представленного в выпускной квалификационной работе, дополняют информацией о семантическом и синтаксическом уровнях языка базу данных, на которой основывается функционирование алгоритма.

Для достижения обозначенной цели были поставлены следующие **задачи**:

- обобщить имеющиеся в научной литературе сведения о семантических и синтаксических валентностях, а также о семантических ролях и об их основных классификациях;
- сформировать на основе этих сведений метод описания семантических ролей, адекватный практическим целям работы, т.е. пригодный для последующего использования в алгоритме и апробации на естественном языковом материале;

- проанализировать материал русской устной речи с опорой на составленное описание, в случаях необходимости – дополнить или модифицировать описание;
- выявить на основе проанализированного материала наиболее часто активирующиеся семантические роли в клаузах с глаголом *быть*, а также типичные способы их выражения в русской устной речи.

Решение перечисленных задач было осуществлено с помощью следующих **методов**:

- критический анализ научной и методической литературы;
- отбор корпусного материала;
- семантический и синтаксический анализ материала;
- количественный анализ результатов.

Объектом исследования являются контексты употребления глагола *быть* в русской устной речи; **предмет исследования** – реализация семантических ролей глагола *быть* в русской устной речи.

Основным источником **материала** исследования послужили данные Корпуса русской устной речи (<http://russpeech.spbu.ru/>); вспомогательным источником материала исследования был пользовательский устный подкорпус, входящий в состав Национального корпуса русского языка (далее – НКРЯ; <http://ruscorpora.ru/>).

Теоретическая значимость исследования обусловлена вкладом, который она вносит в изучение проблематики семантических ролей и их кодирования. **Практическая значимость** работы обусловлена тем, что ее результаты будут учтены в алгоритме, имитирующем механизм восприятия речи человеком; кроме того, составленная классификация семантических валентностей может быть использована при обучении русскому языку как иностранному.

Глава 1. Теоретические предпосылки исследования

1.1. О понятии «валентность»

Синтаксическая сочетаемость слов описывается как пространство, где взаимодействуют словарь и грамматика языка. Их взаимодействие основано, в том числе, на связи между синтаксической и семантической сочетаемостью языковых единиц [Тестелец 2001: 156].

Так, структуру предложения определяют грамматические признаки входящих в предложение лексических единиц и, в частности, их селективные, или сочетаемостные, свойства. Селективными признаками единицы языка называется ее способность или неспособность сочетаться с другими единицами в одной осмысленной и грамматически приемлемой речевой цепочке [Тестелец 2001: 81]. Грамматическая правильность предложения невозможна, если селективные признаки его лексических составляющих не согласованы друг с другом. *Валентностями* называют и семантические, и синтаксические селективные признаки лексем.

Этот термин был заимствован в лингвистические исследования из химии по очевидной аналогии со способностью атомов образовывать определённые связи с другими элементами [Кацнельсон 1987: 31; Тестелец 2001: 156]. В числе первых влиятельных работ, обсуждающих валентности в лингвистическом смысле этого термина, называют [de Groot 1949] и [Tesnière 1959]¹, в отечественной лингвистике – [Кацнельсон 1948] (см. об этом [Ярцева 1990; Богуславский 1996: 12; Bisang et al. 2007]).

Как и в случае с многими другими лингвистическими терминами, использование понятия валентности серьезно варьировалось как в синхроническом, так и в диахроническом плане, в результате чего на настоящий момент существует целый спектр теоретических контекстов его употребления: от референции к управлению в традиционных западных грамматических описаниях до модели управления в ее генеративистском

¹ В русском переводе – [Теньер 1988].

понимании (см., например, [Chomsky 1965]). Не пытаясь углубиться в историю вопроса, ограничимся следующим примером: тогда как в западноевропейской академической традиции явление валентности изначально рассматривалось как относящееся к определенному семантико-синтаксическому свойству именно и только глагола (см., например, эксплицитную вербоцентричность [Tesnière 1959] и [Mosel 1984] как пример более поздней работы, написанной в рамках этого же подхода), в отечественной традиции (и в большей части современных теоретических работ) под валентностью понимается селективный признак слов вне зависимости от их частеречной принадлежности [Кацнельсон 1948; Haspelmath, Müller-Bardey 1991; Апресян 1995; Matthews 2007].

Итак, можно говорить о многозначности термина *валентность* в академическом дискурсе (как будет показано далее, в этой области есть смежная проблема наличия нескольких наименований для референции к одним и тем же явлениям). Кроме того, даже понимаемое в рамках какого-либо одного конкретного подхода, само явление валентности неизбежно затрагивает такое основополагающее свойство языка, как способность языковых единиц к связи между собой: так, заполнение валентностей было описано как «основной способ объединения значений слов в более крупные смысловые фрагменты» [Богуславский 1996: 12]. Очевидным образом, это понятие оказывается релевантным для огромного пласта теоретических проблем и практических задач, которые пытаются решить исследователи в самых разных областях лингвистики.

Кроме традиционных исследований в области семантики и синтаксиса, валентность выступает как предмет диахронических [Korhonen 1995] и сопоставительных [Bisang et al. 2007: 217–87] исследований, а также исследований ментального лексикона [Emons 2007; Риехакайнен 2015] и т.д.

С появлением компьютерной лингвистики и началом попыток автоматизировать такие процессы, как перевод, распознавание речи и т.д., валентность стала также активно рассматриваться в исследованиях,

посвященных разработке механизмов автоматизации. Так, например, в работах о модели «Смысл – Текст» указывается на то, что сведения о валентностях слова входят в его *модель управления* [Мельчук 1974].

То же самое можно сказать и о других актуальных для данной работы проблемах: так, например, аппарат семантических ролей лежит в основе классификации базовых конструкций А. Е. Кибрика, в рамках которой языки разделяются на четыре типа в зависимости от способа кодирования агенса и пациенса при переходном и непереходном глаголе [Климов 1977; Кибрик 1992, 1997]; само понятие *семантическая роль* приблизительно соотносится, например, с понятием семантических функций в рамках функциональной грамматики [Dik 1978] и D-структурой в генеративистском подходе [Chomsky 1981].

Таким образом, даже поставив перед собой узкие и конкретные прикладные задачи, мы оказываемся вынуждены взаимодействовать с целым полем понятий, связанных с большим количеством теоретических вопросов, рассмотрение каждого из которых требует отдельного исследования². В данной работе, естественным образом, не делается попытки их разрешить, а только рассматриваются релевантные для анализа материала термины и определяется, что под ними будет пониматься в рамках этого исследования.

1.2. Семантическая валентность, предикат, семантический актант

Описания значений многих слов (как в бытовой речи³, так и в виде словарного семантического толкования) часто содержат указание на наличие некоего неопределенного элемента, обязательно участвующего в описываемой ситуации, но не называемого конкретно. Как правило, при референции к такому участнику используются неопределенные местоимения. Например, в

² См. краткое и очень точное описание проблемы в [Mel'c'uk 2004: 1] "... the concepts and the terms used are vague, the same concepts are designated by different terms and the same term is used for different concepts".

³ См., пример такого толкования в повседневной речи: ««Дезавуировать» — это когда, например, кто-то сделал заявление, а потом его начальство это заявление опровергает» [Тестелец 2001: 157].

[MAC] приводится следующее определение семантики глагола *довезти*: «везя, доставить до какого-л. места».

Описываемая глаголом ситуация подразумевает наличие как минимум трех участников: во-первых, перемещаемого объекта X, во-вторых, субъекта Y, осуществляющего перемещение, и/или инструмента, с помощью которого осуществляется перемещение (в случае *довезти* таким инструментом, вероятнее всего, будет некий вид транспорта), и, в-третьих, конечной точки Z, в которую доставляется объект. Тогда как первые два участника входят в значение глагола *везти*⁴, третий эксплицитно выражен в приведённом толковании с помощью именной группы «до какого-л. места».

Исключение любого из них неизбежно приведет к изменению семантики всей лексемы: например, при изъятии третьего участника – финальной точки движения субъекта и/или инструмента – значение глагола *довезти* оказывается близким к значениям глагола *везти*³ и *отвезти*⁵.

Показательный пример анализа валентностной структуры лексемы, описывающей более сложно устроенную ситуацию, приводится в [Апресян 1995: 120], где глагол *арендовать* представлен следующим образом: «А арендует С значит, <...> что за какое-то вознаграждение D лицо А приобретает у другого лица В право на эксплуатацию недвижимой собственности С в течение времени Т». [Апресян 1995: 120]. Соответственно, обязательными участниками ситуации аренды оказываются: тот, кто арендует; то, что арендуют; тот, у кого арендуют; плата, за которую арендуют; и срок, на который арендуют. Исключение срока как участника ситуации превращает акт аренды в акт купли–продажи; исключение сдаваемой собственности – в ситуацию, близкую к займу и т.д. [Там же] (см. также другие подобные

⁴ 1. Передвигать, перемещать, заставлять перемещаться какие-л. средства передвижения, а также перемещать что-л. при помощи средств передвижения. || Передвигаясь, перемещаясь, доставлять кого-, что-л. куда-л. (о средствах передвижения) || Иметь при себе, с собой, отправляясь в поездку с целью доставки куда-л., кому-л. 2. (сов. повезти) безл., кому. Разг. Об удаче в чем-л. [MAC].

⁵ 1. Везя, доставить куда-л.; свезти. 2. Увезти в сторону, на некоторое расстояние от кого-, чего-л. [MAC].

примеры: *ехать*, *писать* и др. [Там же: 122–23, 134]; *продавать* [Тестелец 2001: 159]; *наказать*, *наказание* [Мельчук 1974: 134]).

Итак, несмотря на то что лица или предметы, к которым отсылают языковые выражения, являющиеся участниками ситуации, меняются в зависимости от контекста конкретного употребления глагола, исчерпывающее описание семантики этих (как и многих других) лексем невозможно без указания на сам факт обязательной инкорпорации участников в значение лексемы.

Следует также отметить, что в осмысленном предложении такие элементы, хотя и «неопределенные или до конца не проясненные» [Тестелец 2001: 157], не могут быть реализованы с помощью указания на любое случайное лицо или предмет. Точно так же, как значение всей лексемы «диктуется» наличием или отсутствием определенного участника, сам участник должен соответствовать определенным семантическим критериям, заданным значением уже лексемы. Так, например, в обсуждаемом случае глагола *довезти* объект действия должен быть перемещаем, субъект и/или инструмент должен обладать способностью его переместить, третий участник должен реферировать к значению места, а не, например, времени, и т.д.

Таким образом, значения многих лексем включают в себя обязательные переменные единицы с некоторыми заданными семантическими параметрами. По определению Я. Г. Тестельца (которое будет приниматься за опорное в данной работе), любая такая «(несвязанная) переменная X, входящая в толкование (описание значения)» лексемы L, называется **семантической валентностью** лексемы L [Тестелец 2001: 158].

Здесь мы впервые непосредственно сталкиваемся с описанным в п. 1.1. свойством исследований и академических дискуссий, посвященных семантическим и синтаксическим валентностям и смежным темам, а именно с большим разнообразием терминологии, используемой в этой области для описания часто очень близких или одних и тех же явлений и, наоборот, с многозначностью терминов. Эта проблема актуальна при рассмотрении как

мировой, так и только отечественной научной литературы. Поэтому, несмотря на то что создание еще одного ряда новых определений (или еще одной новой классификации) для описания обсуждаемых явлений лежит за пределами задач данной работы, представляется важным перечислить основные принятые варианты референции ко всем интересующим нас явлениям.

Так, синонимичным семантической валентности термином является **партиципант**⁶; в свою очередь, каждая лексема, обладающая хотя бы одной семантической валентностью, может называться **предикатным словом**, **предикатом** [Тестелец 2001: 158], а также **валентным словом** [Богуславский 1996: 23].

Языковые выражения, называющие обязательных участников описываемой ситуации, т. е. те, которыми заполняются валентности предиката в предложении, называются его **семантическими актантами** [Тестелец 2001: 160]. Для обозначения отношений между семантической валентностью и заполняющим ее актантом, используют в том числе слово *соответствие* [Тестелец 2001: 160]. Снова обращаясь к глаголу *довезти*, можно сказать, что партиципант Z соответствует семантическому актанту «пропасть» в предложении «Поезд довез его до пропасти».

Семантические актанты должны быть «достаточны и необходимы» [Апресян 1995: 120], т. е. исчерпывающе и неизбыточно описывать именно называемую лексемой ситуацию. Следует отметить, однако, что, хотя этот критерий в целом удовлетворяет требованиям данного исследования, он оказывается достаточным далеко не во всех случаях. Так, если отталкиваться только от него, возникает следующая проблема: многие предикаты описывают ситуации, подразумевающие большое количество неисключаемых участников. Например, *родиться* или *умереть* можно только в каком-либо конкретном месте. Тем не менее, интуитивно представляется неправильным говорить о том, что в ситуациях смерти и рождения участник места имеет

⁶ См. также обсуждение понятия содержательной валентности в [Кацнельсон 1987].

такой же статус, как, например, в ситуациях, описанных глаголами движения (*уйти, отправиться* и т.д.) [Тестелец 2001: 169].

Таким образом, необходим какой-либо дополнительный критерий или механизм идентификации семантических валентностей. В качестве решения этой проблемы предлагались разные подходы. Например, т.н. «тест Паневой», или «тест диалога», основан на принципе **облигаторности** семантической валентности, подразумевающим, что партиципантами должны быть признаны только те участники ситуации, которые не могут быть неизвестны говорящему [Panevova 1978; 1994]. Другой способ, предложенный в [Тестелец 2001: 171–172], исходит из предположения о том, что семантические валентности обязательно оказывают влияние на грамматическую структуру предложения. Оба эти способа имеют свои недостатки: в первом случае, невозможно учесть все обстоятельства (в т.ч. экстралингвистические), которые могут повлиять на наличие или отсутствие доступа говорящего к информации об описываемой ситуации. Кроме того, тест нельзя использовать для анализа абсолютных конструкций⁷. Во втором случае из семантических валентностей оказываются исключены нереализуемые валентности (см. п. 1.2.3.) [Тестелец 2001: 170–172].

1.3. Синтаксическая валентность, актант, сирконстант, распространитель

Синтаксической валентностью называют селективный признак, согласно которому «лексема L может иметь в качестве вершины или в качестве зависимого слово W (или составляющую фразовой категории C)»⁸ [Тестелец 2001: 162]. Синтаксические валентности классифицируют в зависимости от статуса W или C: валентность, определяющую, какие зависимые единицы может присоединять L, называют *активной* (или *внутренней*); синтаксическую валентность, указывающую, к какой вершине L может быть

⁷ Подробнее об абсолютных конструкциях см. [Апресян 2000].

⁸ См. также обсуждение понятия формальной валентности в [Кацнельсон 1987] и, например, [Mosel 1984] как вариант альтернативной классификации.

присоединена сама в качестве зависимого, называют *пассивной* (или внешней) [Там же]. В Табл. 1 приведены примеры активных и пассивных валентностей из [Тестелец 2001: 163].

Таблица 1. Примеры активных и пассивных валентностей

	Соответствует партиципantu	Не соответствует партиципantu
Активная валентность	<i>Лучше меня</i> <i>Иван уходит</i> <i>Вид на</i>	<i>Пишет хорошо</i> <i>Очень красивый</i> <i>Разговаривают в</i>
Пассивная валентность	<i>Пишет хорошо</i> <i>Красивый конверт</i> <i>Источник шума</i>	<i>Иван уходит</i> <i>Разговаривают в</i> <i>Что уходит</i>

Природа активных и пассивных синтаксических валентностей неоднородна: набор пассивных синтаксических валентностей лексемы преимущественно обусловлен ее частеречной принадлежностью, тогда как активные синтаксические валентности слова во многом определяются его лексическим значением (см., например, сравнение пассивных синтаксических валентностей лексем *убил*, *убийство* и *убитый* и активных синтаксических валентностей лексем *умер* и *убил* в [Тестелец 2001]).

Соответственно, для некоторых групп лексем небезосновательно делаются обобщения, касающиеся типичной для них роли зависимого или вершины. Так, например, в полной речевой цепочке все предлоги обязательно занимают позицию зависимого; глаголы и предлоги склонны выступать как вершина и т.д. Такие, свойственные конкретным частям речи, синтаксические валентности называют **категориальными**; в тех случаях, когда синтаксическая валентность свойственна определенной подгруппе внутри класса частей речи, ее называют **субкатегориальной** (например, группа предлогов, входящих в состав именной группы в конкретном падеже) [Тестелец 2001: 163].

Синтаксические валентности также подразделяют на **обязательные** и **необязательные**, или **факультативные**, по признаку наличия или отсутствия влияния на грамматическую правильность предложения. Так, незаполнение обязательной валентности актантом влечет за собой грамматическую неправильность предложения, тогда как незаполнение необязательной валентности допускается языком (подробнее см. [Апресян 1995: 149; Тестелец 2001: 173–175]).

Языковые выражения, присоединяющие или присоединяемые словом для заполнения его активной или пассивной синтаксической валентности, могут как соответствовать, так и не соответствовать ее семантическим валентностям. Для обозначения единиц, заполняющих активные синтаксические валентности лексемы L и при этом соответствующих некоторой ее семантической валентности, используют термин **синтаксический актант**⁹; для обозначения единиц, заполняющих активные синтаксические валентности лексемы, но при этом не соответствующих никакой ее семантической валентности, используют термин **сирконстант** [Тестелец 2001: 166].

В качестве иллюстрации можно привести цитату Ю. Д. Апресяна об обсуждавшемся выше глаголе *арендовать*: «Ничто в ситуации аренды не требует указания того, по какой причине, где, когда, с какой целью она осуществилась, хотя <...> словоформы с причинным, местным, временным и целевым значениями <...> вполне присоединимы: арендовать из-за безземелья охотничьи угодья, арендовать прошлым летом под Москвой садовый участок, арендовать клуб для проведения собрания» [Апресян 1974: 120]. Соответственно, в перечисленных примерах употребления глагола *арендовать* зависимые от него языковые единицы можно разделить на актанты и сирконстанты следующим образом (см. Табл. 2).

⁹ См. также выделение semantic, deep-syntactic и surface-syntactic актантов в [Mel'cuk 2004].

Таблица 2. Примеры актантов и сирконстантов глагола «арендовать»

	Актант	Сирконстант(ы)
1	охотничьи угодья	из-за безземелья
2	садовый участок	прошлым летом под Москвой
3	клуб	для проведения собрания

Подход к разделению и актантов и сирконстантов (также *complements* и *adjuncts* в англоязычной традиции), изложенный в приведённых определениях (см. также труды московской семантической школы, например [Жолковский 1964; Богуславский 2020], и работы в рамках модели «Смысл – Текст», например [Мельчук 1974: 85–86]), полностью удовлетворяет прикладным задачам данного исследования, и именно им мы будем руководствоваться далее. Однако важно сказать, что, как и другие интересующие нас понятия, актанты и сирконстанты были и продолжают быть предметом развернутой дискуссии: см., например, [Somers 1987: 12–28; Bonami 1999; Храковский 1998; Богуславский 2020].

Продолжая разговор о проблеме терминологии, следует также упомянуть, что в научной литературе термины *валентность* и *актант* часто по умолчанию означают *синтаксические* валентность и актант [Тестелец 2001]. Вероятно, здесь прослеживается влияние уже упоминавшейся ранее работы [Tesniere 1959], в которой термины *актант* и *сирконстант* были впервые введены для описания именно синтаксического уровня языка. В случае, когда подразумеваются *семантические* валентность и актант, они обычно эксплицитно называются с помощью соответствующего прилагательного.

Кроме того, в литературе часто используется термин *аргумент* / *argument* [Плунгян 2000: 164] (в англоязычной традиции также *term*, см., например, [Van Valin 2001]), который может отсылать как к семантическому, так и к синтаксическому актанту (см., например, [Godard 1996; Whaley 1997] и упоминание первой работы в [Mel'cuk 2004: 6]). В тексте

данной работы этот термин не будет использоваться, чтобы избежать нежелательной ситуации многозначности.

Еще одно относительно близкое понятие – *распространитель* – используется уже в русскоязычной литературе для обозначения «словоформы, связанной с глаголом тем или иным видом подчинительной связи» [Русская грамматика 1980].

Завершая тему синтаксических валентностей, следует обозначить отношения между ними и партиципантами. В целом, их можно охарактеризовать следующей цитатой Я. Г. Тестельца: «Важнейшая особенность человеческого языка заключается в том, что семантическая и синтаксическая сочетаемость слов в большой степени согласованы» [Тестелец 2001: 163]. Итак, лексема с определенной семантической валентностью будет склонна иметь и синтаксическую валентность, реализованную соответствующим семантической валентности синтаксическим актантом.

1.4. Некоторые отклонения от прототипической ситуации

Следует отметить, что некоторые из перечисленных в предыдущих пунктах фактов являются своего рода идеалистическими обобщениями, которые отображают нормальные, но не единственно возможные в языке ситуации. Далее рассмотрим несколько исключений из описанных общих правил, представляющих наибольший интерес с точки зрения работы с устной речью.

Так, говоря об обязательности или необязательности валентностей, нельзя не упомянуть, что в языке допускается возможность незаполнения почти любой валентности в результате **эллипсиса**, или **эллиптического сокращения**: т. е. «сокращения повторяющейся части в одном или нескольких предложениях в связном тексте или в диалоге» [Тестелец 2001: 172], например [Там же]:

- (1) *Иван вошел и Иван поздоровался — Иван вошел и поздоровался;*

(2) *Этот тюльпан расцветает раньше других тюльпанов — Этот тюльпан расцветает раньше других;*

(3) *Саша купил гвозди! — Да, Саша купил гвозди — Да, купил.*

В результате эллипсиса может опускаться любой элемент предложения; это является верным даже для языков, в которых заполнение определенных валентностных «слотов» (например, подлежащего в английском языке) является обязательным [Там же].

Еще одним отклонением от типичной ситуации обязательности / необязательности валентностей является существование **обязательных сирконстантов**. Сирконстанты преимущественно факультативны; однако в ряде случаев можно говорить об их контекстно обусловленной обязательности (4), (5) [Тестелец 2001: 181].

(4) *У него *{серые} глаза;*

(5) *Гранаты *{не той} системы.*

Как следует из примеров, есть основания предполагать, что контекстная обязательность сирконстантов при предикате *быть* будет наблюдаться гораздо чаще, чем при среднем глаголе русского языка.

Смежной проблемой является существование **нереализуемых** семантических валентностей, т. е. таких, которым не соответствует никакая синтаксическая валентность. Заполнение нереализуемой валентности делает предложение грамматически неправильным: так, в примере (6) разрешается реализация только одного из трех–четырёх возможных партиципентов (кто, чем, куда, посредством чего):

(6) *Он промахнулся (*кулаком* мне в лицо /*в окно / *мимо окна бутылкой)* ([Мельчук 1974: 134], цит. по [Тестелец 2001: 167]).

И, наоборот, в языке возможны случаи, когда синтаксические валентности и заполняющие их актанты не соответствуют семантическим [Тестелец 2001: 168]. Они могут заполняться особыми языковыми единицами, имеющими форму (означающее), но не имеющими значения. Это, например, формальные подлежащие *it* и *there* в английском языке в конструкциях типа

англ. *It is raining*, где лексема (в данном случае, *rain*) не предполагает наличия никаких участников ситуации¹⁰.

Еще один комментарий должен быть сделан о явлении **сентенциального актанта**, т.е. такого актанта, который соответствует элементарному предложению [Тестелец 2001: 180]. Выше для описания способов реализации валентностей использовались обобщающие понятия «языковые выражения» или «языковые единицы». Необходимо уточнить, что в качестве таких выражений могут выступать не только отдельные словоформы или даже предложные (7) и именные (8) группы, но и целые придаточные предложения (9).

(7) *Знаю о его приходе*;

(8) *Твое присутствие необходимо*;

(9) *Знаю, что он пришел*¹¹.

1.5. Семантическая роль, классификации семантических ролей

1.5.1. Понятие и общие принципы выделения семантических ролей

В [Tesniere 1959] приводится часто цитируемая (см., например, [Жолковский 1964: 3; Mel'cuk 2004: 1 и др.]) метафора глагольного узла как «маленькой драмы», в которой обязательно происходит некоторое действие, а также, как правило, фигурируют участвующие в нем «действующие лица» и какие-либо обстоятельства происходящего. Соответственно, актанты можно описать как выражающие «участников» действия, а сирконстанты – как «обстоятельства, в которых разворачивается действие» [Теньер 1988]. Если распространить эту метафору, во-первых, на все предикаты и, во-вторых, на семантический план языка, можно заметить, что «действующие лица» разных «драм» часто исполняют одну и ту же функцию; общими свойствами также обладают и «обстоятельства».

Так, в примерах (10) и (11) с помощью разных грамматических структур описываются два совершенно разных события, в которых участвовали разные

¹⁰ См. больше специфичных для английского языка случаев и проблем в [Jackendoff 1990].

¹¹ Примеры из [Тестелец 2001: 180]

люди при разных обстоятельствах. Тем не менее, очевидно, что Вася и племя каннибалов имеют в приведенных предложениях общую «роль»; то же самое можно сказать и о Леше и капитане, и об указании на время реализации ситуации.

(10) **Вчера Леша был избит Васей.**

(11) **Племя каннибалов расчленило капитана в 1530 году.**

Общность участников ситуации внутри перечисленных пар интуитивно заметна, несмотря на то что в «Леша» выражен подлежащим, а «капитан» – дополнением; актант «племя каннибалов» указывает на группу людей, а «Вася» – конкретное лицо; «вчера» является наречием, а «в 1530 году» – именной группой, состоящей из существительного, количественного числительного и предлога и т.д. Позиция участников в предложении различна также с точки зрения порядка слов.

И наоборот, в случаях, показанных в примерах 12–14, которые взяты из [Comrie 1989: 58], участники *John*, *door* и *key* по очереди «становятся» подлежащим при лексически одном и том же глаголе; тем не менее очевидно, что во всех трех предложениях за каждым из них закреплена одна и та же функция.

(12) *John opened the door with the key.*

(13) *The key opened the door.*

(14) *The door opened.*

Итак, необходим некий аппарат, который позволил бы, во-первых, описать сходство Леша, капитана и т.д., несмотря на отсутствие какого-либо структурного единообразия в плане их выражения, и, во-вторых, разграничить участников, сходных с точки зрения классификации членов предложения (или *grammatical relations* в англоязычной традиции), но «играющих» интуитивно разную «роль».

Для решения этой проблемы выделяют явление, для референции к которому в рамках разных подходов используют целый ряд терминов (например, *тематическая роль*, *тета-роль*, просто *роль*, *семантическая*

функция, глубинный падеж, глубинная структура и т.д. [Lyons 1966, 1967; Филлмор 1968; Jackendoff 1990; Апресян 1995; Плунгян 2000; Тестелец 2001]), из которых мы остановимся на понятии **семантической роли**, предложенном Ч. Филлмором (см., например, [Филлмор 1968]). Семантическая роль актанта «отражает общие свойства участников определенных групп ситуаций» и является частью значения присоединяющего его предиката [Апресян 1995: 120; Плунгян 2000: 164].

Между семантическими ролями и актантами, являющимися их планом выражения, нельзя установить точное соответствие [Апресян 1995: 26]. Однако важно отметить, что плоскость семантических ролей не может быть полностью отчуждена от синтаксической структуры предложения. Так, Б. Комри описывает их связь следующим образом: «Для того чтобы выделение семантической роли принималось на уровне универсального инвентаря семантических ролей, хотя бы один язык должен иметь некий грамматический коррелят семантического различия, на котором основано это выделение»¹² [Comrie 1981: 53].

Не ставя перед собой цель обсудить все аспекты этого взаимодействия, приведем следующий пример: правило о том, что «у одного предикатного слова не может быть двух актантов с одной и той же семантической ролью» [Fillmore 1969: 117; Тестелец 2001: 209]¹³, коррелирует с запретом в языке на сочинённое выражение двух участников, семантические роли которых различаются [Там же].

Это же свойство семантических ролей является важным критерием, используемым в проблеме разработки их классификации, т.к. из него следует, что только актанты с одинаковой семантической ролью могут быть сочинены в предложении. Нужно отметить, что у этого общего принципа есть ряд

¹² “In order for a distinction of semantic role to figure in a universal inventory of semantic roles, it must be the case that at least one possible language has some grammatical correlate of this semantic distinction”.

¹³ См. также обсуждение на материале английского языка и специфические для английской грамматики примеры в [Jackendoff 1990: 59-61]; о материале русского языка см. [Муравенко 1998].

исключений. Так, актанты и сирконстанты могут сочиняться несмотря на несовпадение их семантических ролей в следующих случаях:

- если они выражены кванторными или вопросительными местоимениями: *Нигде и никогда не бывало и не может быть материи без движения, движения без материи...* [В. И. Ленин. Карл Маркс (1913)];
- если они содержат указательные местоимения: *...спрашивается, что еще нужно для полного счастья в такой день и в такой компании?* [Вильям Смит. Американец в России — оказывается это очень просто и даже интересно! // «Знание-сила», 2014];
- если в их состав входит одна и та же лексема: *...говорю с лингвистом и о лингвисте* [Санников 1989: 15–25] цит. по [Тестелец 2001: 212];
- если содержащее их предложение включает в себя двойное отрицание, выраженное при каждом из сочиняющихся актантов / сирконстантов: *Не он, и не ко мне приходит.* [С. В. Соловьев. Барка // «Волга», 2012]¹⁴.

Кроме того, актанты или сирконстанты с одинаковыми ролями не всегда могут сочиняться: например, при осуществляемой ими обоими референции к месту реализации ситуации сочинение возможно, только если два обозначаемых места не пересекаются [Тестелец 2001: 212]:

(15) **В России и в Петербурге.*

Тем не менее, такая закономерность предоставляет в распоряжение исследователей в этой области эффективный в большинстве случаев тест. Так, излишняя детализация семантики предполагаемых ролей приводит к тому, что не запрещаются даже такие предикаты, которых на самом деле язык не допускает (см. обсуждение глаголов со значениями еды и питья в [Тестелец 2001: 210]). И наоборот, в случае если семантика разных ролей формулируется слишком обобщенно, многие глаголы естественного языка будут

¹⁴ Классификация из [Тестелец 2001: 212]; все примеры, кроме того, на который дана ссылка в тексте, взяты из [НКРЯ].

«запрещены». Так, если выделить условную семантическую роль Предмет 1 и классифицировать каждый неодушевленный предмет, который может каким-либо образом передвигаться в пространстве, как относящийся к ней, не будет «разрешен» глагол *довезти* (16).

(16) *Он довоз чемодан* (Предмет 1) *поездом* (Предмет 1).

Кроме приведенных выше «синтаксических» критериев, при выделении семантических ролей как правило исходят из следующих более общих предпосылок:

- семантические роли универсальны для все всех языков [Comrie 1981; Van Valin 1993: 50], и, соответственно, предполагается, что их состав и количество в рамках предлагаемого инвентаря должны делать этот инвентарь пригодным для применения на любом языковом материале;
- классификация семантических ролей не может быть излишне дробной: наименования семантических ролей должны описывать «в значительной степени обобщенные» группы языковых единиц [Плунгян 2000: 164].

Как несложно заметить, даже принимаемые в совокупности, эти критерии предоставляют большую свободу при разработке инвентаря семантических ролей. Действительно, только количество выделяемых ролей в рамках разных перечней варьируется от (приблизительно) 10 ролей¹⁵ (например, семь в [Гак 1969], девять в [Fillmore 1971]) до 50-ти [Леонтьева 1968]; уже из этих цифр очевидно, что не может не варьироваться и содержание.

Большинство авторов более или менее единообразно выделяют самые частотные из семантических ролей и определяют их основные черты; однако со снижающейся частотностью ролей вариативность их трактовок значительно возрастает [Тестелец 2001: 214]. Меняются и перечни

¹⁵ Здесь не учитываются классификации гипер- или макроролей, некоторые из которых могут включать в себя только две роли: например, протоАгенс и протоПациенс в [Van Valin 1993].

семантических ролей в рамках подхода одного исследователя (или одной группы исследователей): см., например [Fillmore 1968; Fillmore 1977] с говорящими названиями *The case for case* и *The case for case reopened* соответственно, в которых можно проследить эволюцию взглядов автора и, в частности, переход к более сложному и детализированному инвентарю семантических ролей. Кроме того, даже достаточно частотные и устойчиво выделяемые в среднем семантические роли, такие как бенефактив, реципиент, адресат (см. след. пункт), часто совмещаются под одним из наименований, что вносит дополнительные сложности в процесс классификации [Тестелец 2001: 213–14].

Таким образом, составление некой базы данных, которая бы включала в себя абсолютно все существующие наименования для всех выделяемых семантических ролей в рамках всех существующих подходов, представляется достаточно бесполезным в контексте прикладных задач данного исследования. Не пытаясь задокументировать все случаи корреляции (или ее отсутствия) между различными перечнями, в следующем пункте приведем список семантических ролей (как «центральных», так и «периферийных»), которые целесообразно выделять для проведения анализа русской устной речи.

Этот список, созданный пока только на основе существующих в литературе классификаций, будет опорным при создании классификации семантических ролей актантов и сирконстантов глагола *быть*.

1.5.2. Различные инвентари семантических ролей

На основе [Fillmore 1968; 1969; 1977¹⁶; Гак 1969; Jackendoff 1990; Апресян 1995; Плунгян 2000; Тестелец 2001; Риехакайнен 2015] был составлен обобщающий список распространенных (как в существующих классификациях, так и, предположительно, в речи, которую эти классификации призваны отображать) семантических ролей.

¹⁶ Перевод наименований ролей приводятся вслед за [Апресян 1995].

Следует отметить два факта о перечисленных выше источниках: во-первых, классификации в [Плунгян 2000; Тестелец 2001] тоже являются результатами обобщения; во-вторых, список, приведенный в [Апресян 1995: 125–126], охарактеризован в работе-первоисточнике как классификация семантических валентностей, а не семантических ролей. Однако референция к партиципантам по их семантическим ролям не невозможна и в целом достаточно приемлема (см. [Крылов и др. 2004] со ссылкой на [Мельчук, Холодович 1970]).

Значительная часть спорных случаев, возникающих при попытках обобщить имеющиеся классификации, связана не только с разными подходами к выделению семантических ролей, но и с частым пересечением присваиваемых им наименований. Так, например, название *Цель* используется в разных классификациях для обозначения как 1) одной из семантических ролей места, так и, по определению [МАС]¹⁷, 2) «Того, к чему стремятся, чего хотят достичь». Для разрешения этой проблемы был принят принцип под условным названием *критерий минимальной полисемии*, который можно сформулировать следующим образом:

➤ из нескольких «соперничающих» вариантов наименования выбирается тот, который наименее склонен быть использованным для обозначения другой семантической роли.

Т.е. для обозначения конечной точки движения (1) принималось другое название, а наименование *Цель* присваивалось семантической роли (2).

В дальнейшей работе с корпусом предполагается опираться на следующий базовый список семантических ролей:

▪ Агенс: активный участник ситуации, одушевленный и волитивный. Иницирует действие, контролирует его (т. е. может его и прекратить), в процессе осуществления действия затрачивает свою

¹⁷ Значение ряда семантических ролей интуитивно понятно носителю языка, но редко формулируется эксплицитно. Здесь и далее словарные толкования [МАС] использовались как основа для формулировки характеристик таких семантических ролей.

энергию;

- Пациент: пассивный участник ситуации, контролем над которой он не обладает. Подвергается серьезным изменениям в процессе осуществления действия другим(и) участником(ами), таким как смена «формы, вида, положения в пространстве» [Тестелец 2001: 213], экзистенциального статуса (существование / несуществование);
 - Экспериментер: участник ситуации, воспринимающий информацию посредством органов чувств или непроизвольно испытывающий некое чувство или эмоцию;
 - Стимул: участник ситуации, являющийся источником информации, воспринимаемой посредством органов чувств; либо причина непроизвольно переживаемого чувства или эмоции;
 - Инструмент: в нормальном случае¹⁸ неодушевленный участник ситуации, используемый для осуществления действия, но не претерпевающий изменений в его процессе или в его результате;
 - Средство: в нормальном случае неодушевленный участник ситуации, используемый для осуществления действия и претерпевающий изменения (обычно расходуемый) в его процессе;
 - Адресат: участник ситуации, которому агент направляет информацию;
 - Реципиент (получатель): участник ситуации, который становится обладателем пациента в результате действия агента;
 - Бенефактив: участник ситуации, интересы которого оказываются затронуты в процессе ее осуществления, либо в ее результате, но не являющийся ни агентом, ни пациентом этой ситуации;
 - Контрагент: участник ситуации, против которого направлено действие, осуществляемое агентом;

¹⁸ Нельзя исключить возможность (пусть и маловероятной) ситуации, в которой инструмент будет одушевлен, например: *ребенок ударил брата черепахой*.

- Начальная точка движения: место, из которого направлено движение;
- Место: пространство, в котором локализована ситуация;
- Направление: «место или сторона, в которую устремлено движение» [MAC];
- Маршрут: путь перемещения участника ситуации;
- Источник: «то, откуда берется» материальный или нематериальный ресурс (предмет, информация и т.д.) [MAC];
- Аспект: область проявления или применения свойства или действия;
- Причина: участник ситуации, «непосредственно порождающий, обуславливающий» описываемую ситуацию, либо «основание, повод, предлог» ее осуществления [MAC];
- Время: локализация ситуации на временной оси;
- Срок: период времени, в течение которого реализуется описываемая ситуация;
- Дистанция: расстояние, на котором осуществляется действие [MAC];
- Условия: обстоятельства осуществления действия;
- Состояние: «физическое самочувствие или настроение, расположение духа» одушевленного участника ситуации; «положение, в котором находится» неодушевленный участник ситуации [MAC];
- Содержание: «то, о чем повествуется, что излагается, изображается где-л.» [MAC]; также участник ситуации, являющийся объектом мыслительной деятельности;
- Способ: образ осуществления действия [MAC].

Теперь, располагая данным репертуаром, можно сделать несколько наблюдений о типичных проблемах, возникающих при попытках определения семантических ролей актантов и сирконстантов.

1.5.3. Определение семантической роли актантов: типичные проблемы и подходы к их решению

Все существующие классификации включают в себя семантическую роль Агенса / Субъекта (или, по крайней мере, нечто близкое к ней); большинство также выделяют обладание **контролем** над ситуацией как один из основных признаков этого участника. Тем не менее, далеко не всегда можно однозначно определить, действительно ли «кандидат» на роль агенса обладает контролем (17), или какой из нескольких партиципантов является главным «контролером» (18).

(17) *John rolled down the hill* [Comrie 1981: 58]¹⁹.

(18) *Делегация дипломатов купила у хозяев слона.*

В примере (17) Джон может катиться как контролируя ситуацию, так и не контролируя ее; в (18) описывается действие, которое в нормальном случае контролируется обеими сторонами: и продавцом, и покупателем²⁰.

В рамках различных подходов такие случаи неоднозначности решаются разными способами. Наиболее прямолинейный из них – выделение отдельной, более специфической семантической роли.

Например, для референции к ролям типа роли Джона в ситуации (17) в рамках некоторых подходов выделяют семантическую роль Темы. Ее можно приблизительно описать как участника, в меньшей степени контролирующего ситуацию, чем агенса, но и не претерпевающего существенных изменений (в отличие от пациенса) (см., например, [Van Valin 1993: 44]). Выделение такого менее эксплицитно неволитивного участника позволяет, например, отделить участника неоднозначной ситуации типа (17) от более прототипического пациенса в ситуациях типа (19):

(19) *Mary rolled John down the hill* [Comrie 1981: 58].

Другим возможным решением является использование ряда синтаксических и семантических критериев для определения

¹⁹ См. также обсуждение сходных случаев неоднозначности в [Lyons 1967].

²⁰ См. также обсуждение сходных случаев потенциально двуагентивных ситуаций в [Апресян 1995: 27].

рассматриваемого партиципанта в одну из уже существующих ролей. Так, в [Тестелец 2001: 213] приводятся следующие примеры возможных критериев определения главного контролера (ГК) ситуации²¹:

- способность ГК контролировать референцию невыраженного подлежащего целевого оборота;
- в императиве 2-го лица ГК становится адресатом;
- предложение может быть переформулировано в «демиактивную» конструкцию типа *Мне работается*, где позиция имени в дательном падеже занята ГК;
- предикат способен присоединять агентивные обстоятельства: допустимо предложение *Х намеренно спровоцировал Y-а*; следовательно, Х является ГК²².

Однако, как можно заметить, эти критерии нельзя применить ко всем спорным случаям: так, например, (17) допускает присоединение агентивных обстоятельств, что, тем не менее, никак не разрешает изначальную многозначность.

Точно так же, как определенный участник ситуации может не вписываться ни в одну из предложенных в инвентаре ролей, возможна и ситуация, когда две роли «совмещаются» в одном участнике. См., например, (20), (21):

(20) *Он (Источник–Агенс) преподает мне математику* [Апресян 1995: 26].

(21) *Начертить <...> по линейке (Инструмент–Место–Маршрут) на ватмане (Инструмент–Место)* [Муравенко 1998: 80], также обсуждается в [Тестелец 2001: 212].

Здесь, опять же, возможно создание еще одной роли с соответствующим узким значением либо ориентация на какой-либо синтаксический критерий (см., например, краткое изложение подхода Филлмора в [Апресян 1995: 27-

²¹ Первые три критерия были предложены в [Булыгина 1982].

²² Критерий и пример из [Кацнельсон 1972].

28]). В целом, следует отметить, что сама вариативность систем семантических ролей в рамках разных подходов, а также большое количество неоднозначностей, возникающих при анализе языкового материала, указывают на неизбежную условность (в известной степени) выделения семантических ролей. Итак, представляется возможным говорить об области, обозначаемой семантическими ролями, не как о списке строго противопоставленных единиц, а как о континууме, в котором выделяют и наделяют ярлыками некоторые конкретные «точки» [Comrie 1981: 58; Van Valin 1993: 44].

Метод же выделения обычно основывается на конкретной задаче, стоящей перед автором инвентаря [Плунгян 2011: 114]. Это в особенности верно для работ в области машинного обучения и автоматического распознавания семантических ролей, выполненных в рамках проектов FrameNet (см., например, [Thompson et al. 2003; Shen, Lapata 2007; Coppola et al. 2009]), The Proposition Bank [Palmer et al. 2005a; Palmer et al. 2005b; Palmer et al. 2007], Verbalex [Hlaváčková et al. 2006], FrameBank [Кашкин, Ляшевская 2013] и т.д.

В рамках каждого из этих проектов принимается свой подход, адаптированный, по-видимому, и под теоретический контекст, в котором существуют его авторы, и под технические требования, предъявляемые определенным корпусом. Так, например, система FrameNet базируется на т.н. семантике фреймов Ч. Филлмора и, соответственно, идентифицирует семантические роли участников через типовые ситуации, к которым они принадлежат. Родственная ей система FrameBank, основываясь в том числе на исследованиях в рамках подхода Московской семантической школы, отталкивается от лексических конструкций, свойственных индивидуальным предикатам.

В рамках настоящего исследования мы ориентировались, во-первых, на сформулированный выше инвентарь семантических ролей, а во-вторых,

непосредственно на материал русской устной речи, обсуждаемый во второй главе данной работы.

1.6. Специфика глагола *быть*

Согласно данным Нового частотного словаря русской лексики, глагол *быть* является наиболее частотным глаголом в русском языке, обладая частотностью 12160,7 ipm [Ляшевская, Шаров 2009]. Шесть форм глагола *быть* также входят в список 100 наиболее частотных словоформ в НКРЯ: *было* (26-е место), *был* (32-е место), *быть* (55-е место), *были* (59-е место), *была* (60-е место), *будет* (72-е место).

Итак, глагол *быть* широко представлен в русской речи. Важно отметить, что его употребление чрезвычайно неоднородно: так, выделяют от двух (см., например, [Бенвенист 1974]) до 13 [Апресян et al. 2014]) подвидов *быть*, причем одна такая классификация может использовать как грамматические, так и семантические критерии выделения.

В таблице 3 приводится сопоставление двух таких подходов, принятых в [Апресян et al. 2014] и [МАС], которые были выбраны за опорные в рамках данного исследования.

Таблица 3. Классификация подвидов «быть» в [Апресян et al. 2014] и [МАС]

Классификация по [Апресян и др. 2014]	Классификация по [МАС]
<i>быть 1.1 'являться'</i>	1. Существовать иметься
<i>быть 1.2 "быть тождественным"</i>	<i>в значении связки между подлежащим и именным сказуемым</i>
<i>быть 2.1 'находиться'</i>	<i>2. Находиться, присутствовать где-л. // Располагаться, размещаться, находиться.</i>
<i>быть 2.2 'прибывать куда-л.'</i>	<i>3. Приходить, приезжать куда-л., посещать кого-л.</i>
быть 3.1 'иметься'	1. Существовать иметься
<i>быть 3.2 'иметь возраст'</i>	
<i>быть 4.4 'происходить, случаться':</i>	<i>4. Происходить, совершаться, случаться.</i>
<i>быть 5.1 "уверенность в неизбежности события"</i>	

быть 5.2 'уверенность в неизбежности плохого'	
быть 5.3 'надо прекратить ситуацию'	
быть 5.4 'надо прекратить действие'	
<i>быть 6.1 в составе формы будущего времени</i>	<i>в составе формы будущего времени</i>
<i>быть 6.2 в составе формы страдательного залога</i>	<i>в составе формы страдательного залога</i>

Полностью совпадающие типы выделены курсивом, частично совпадающие показаны жирным курсивом (так, в [Апресян и др. 2014] этот подвид выделяется на основе семантического критерия, его коррелят в [МАС] – на основе синтаксического). И, наконец, случай, когда одному подвиду соответствуют два разных, отмечен жирным шрифтом.

Данная таблица иллюстрирует, во-первых, разнообразие мнений, которые встречаются в научной литературе, затрагивающей классификацию случаев употребления *быть*; и, во-вторых, она показывает невозможность полного разграничения «глубинной» семантики и более внешних, формальных критериев при обсуждении этого глагола.

Таким образом, в рамках данного исследования мы будем допускать возможность выделения таких формальных критериев для решения вопросов описания семантических ролей.

Глава 2. Анализ употребления клауз с глаголом *быть* в русской устной речи

2.1. Контекст и прикладные задачи исследования

Как отмечалось во Введении, практическая значимость данной работы обусловлена ее вкладом в разработку алгоритма восприятия речи, приближенного к естественному механизму восприятия речи человеком.

Некоторые «чисто технические» аспекты, сопряженные с описанной выше задачей работы, оказали значительное влияние как на метод исследования в широком смысле, так и на некоторые частные решения, принимаемые в его процессе. Такие случаи оговариваются отдельно в п. 2.4 данной главы, посвященном непосредственно классификации ролей и способов их кодирования, а также адаптации опорного списка для описания клауз с глаголом *быть*.

Кроме того, вхождение данного исследования в другое, более крупное, определило выбор основного источника материала исследования: им стал Корпус русской устной речи (<http://russpeech.spbu.ru/>), разрабатывающийся в рамках этого же проекта. Вспомогательным источником материала был подкорпус устной речи НКРЯ (<https://ruscorpora.ru/new/>).

Это прикладное измерение исследования предполагает последовательное решение следующих задач:

- 1) составление опорного перечня способов кодирования в русской устной речи актантов и сирконстантов, которые могут быть описаны в терминах классификации семантических ролей, предложенной в п. 1.5.2;
- 2) рассмотрение клауз с глаголом *быть* в материале русской устной речи и модификация списка семантических ролей для отражения специфики глагола *быть* на основе наблюдаемых закономерностей;
- 3) составление перечня способов кодирования актантов и сирконстантов глагола *быть* в проанализированном материале;

4) описание особенностей клауз с глаголом *быть* в русской устной речи с т.з. активации в них семантических ролей, а также синтаксических способов их кодирования участников описываемых ситуаций

2.2. Опорный перечень способов кодирования семантических ролей в русской устной речи

В п. 1 данной работы приводился опорный список семантических ролей. Однако, чтобы начать выделять эти семантические роли в предложении естественного языка, необходимо иметь также некий инвентарь наиболее частотных способов синтаксических способов реализации этих ролей. Так как в рамках данной работы в конечном итоге предполагалось применение этого инвентаря к устной речи, было принято решение отталкиваться не от словарных описаний, а от руководств по русскому языку как иностранному. При этом мы исходили из предпосылки, что именно в подобных работах приводится информация о наиболее употребительных в речи синтаксических конструкциях, которые, соответственно, будут представлены и в анализируемых корпусах устной речи. Итак, был составлен список всех перечисленных в пособиях [Глазунова 2005; Ермаченкова 2009] способов синтаксического оформления распространителей, т.е. сочетания (опционально) предлога и падежа члена предложения. Список был ранжирован по алфавиту, после чего – в случае наличия – распространителям были присвоены названия, отсылающие к наиболее распространенным классификациям (Агенс, Аациенс, Экспериенцер и т. д.). Таким образом случаи наиболее типичного употребления падежей в русском языке согласно [Глазунова 2005; Ермаченкова 2009] были соотнесены с инвентарем семантических ролей, сформулированным в п. 2.

Отдельно был составлен список специфических синтаксических конструкций, в которых лексическая или грамматическая составляющая оказывает влияние на выбор падежа или падежей. В него вошли посессивные

конструкции, конструкции с определёнными числительными, наречиями, прилагательными или наречиями в сравнительной степени, а также пассивные конструкции.

В ряде случаев при составлении опорного перечня для обозначения семантической роли, которой обладает определённый распространитель (или группа распространителей), не находилось коррелята из списка, приведенного в п. 2. Для их описания в список были введены следующие дополнительные элементы:

- **Посессор:** *машина брата;*
- **Замещение:** *вместо тетради;*
- **Исключение** *кроме моих друзей, никто не знал об этом;*
- **Совместность** *разговаривать с преподавателем, приехать с портфелем, встретиться с другом* (примеры из [Глазунова 2005; Ермаченкова 2009]).

Таким образом, на этом этапе работы был составлен опорный список распространителей и соответствующих им способов синтаксического оформления, который редактировался и дополнялся впоследствии при работе с корпусным материалом. Полный результат этой части работы представлен в Приложении 1 (падежи и сочетания «предлог + падеж») и Приложении 2 (конструкции).

2.3. Общая информация о материале исследования

2.3.1. Корпус русской устной речи и работа с ним

Корпус русской устной речи (далее – КРУР) включает в себя звучащие тексты, а также их орфографическую и акустико-фонетическую транскрипцию. На момент осуществления исследования корпус содержит звучащие тексты общим объёмом более 22-х тыс. словоупотреблений, созданные с 40-х гг. XX века до настоящего времени. Это профессиональное дикторское чтение, чтение рядовыми носителями языка и монологическая спонтанная речь.

Поскольку в корпусе на данный момент возможен поиск только по словоформам, поиск контекстов, включающих в себя клаузы с глаголом *быть*, осуществлялся вручную. Формировались запросы, содержащие последовательно все формы глагола (*быть, есть, буду, будешь, будете, будет, будут, был, была, было, были, будь, будьте*). Все было получено 238²³ примеров. Тексты в полученном массиве относились к следующим стилям: профессиональное дикторское чтение (9 вхождений), чтение рядовыми носителями языка (30 вхождения), монологическая спонтанная речь (199 вхождений).

Далее велась работа с отобранными контекстами. Из материала исследования были исключены следующие случаи употребления форм глагола *быть*:

- вхождения, где форма глагола *быть* является показателем времени, модальности и / или наклонения при предикативе (например, *если бы можно было быть великодушным за тысячу долларов...*);
- вхождения, где форма глагола *быть* является частью синтетической формы пассива (например, *«ребенок был вовлечен в некую систему воспитания...»*).

Перечисленные случаи исключались, так как функция глагола *быть* очевидным образом сводится в них только к грамматической; семантическое же оформление и все реализованные (или возможные) в клаузе валентности относятся к лексической части рассматриваемого сказуемого. Так, например, в клаузе *если бы можно было быть великодушным за тысячу долларов...* реализуемые семантические валентности под условными наименованиями Определение («великодушным») и Стоимость («за тысячу долларов») относятся именно к части сказуемого, употребленной в форме инфинитива *быть*, тогда как *было* выполняет грамматическую функцию. Соответственно,

²³ Вводные конструкции, такие как *то есть, может быть*, а также случаи омонимии (например, *есть* в значении «употреблять пищу») отсеивались на самом первом этапе сбора материала.

данное вхождение учитывалось в подкорпусе, сформированном для формы *быть*, но исключалось из подкорпуса, составленного для формы *было*.

Итоговый корпус насчитывает 156 вхождений, из них профессиональное дикторское чтение – 3 вхождения, чтение рядовыми носителями языка – 16 вхождений, монологическая спонтанная речь – 137 вхождений.

В этом подкорпусе следует выделить следующие две частично пересекающиеся группы:

- вхождения, где форма глагола *быть* является частью составного именного сказуемого;
- вхождения, где инфинитивная форма глагола *быть* подчинена другому глаголу.

В первом случае именная часть сказуемого выражалась полным или кратким прилагательным (например, *который позволяет ему, ещё раз говорю, быть великодушным*) или существительным (например, *но это должен быть талант, да?*). Во втором случае среди глаголов, присоединяющих инфинитивную форму *быть*, встречались глаголы желания, необходимости, долженствования и т.д. (например, *и все захотят быть детьми, там, генерала Раевского*).

Несмотря на то что данное исследование должно внести вклад в разработку алгоритма, функционирующего именно на материале Корпуса русской устной речи, такой объем материала представлялся недостаточным. Кроме того, КРУР продолжает пополняться и, следовательно, представлялось важным не только описать те контексты, которые представлены в нем сейчас, но и получить более универсальные данные. Для расширения материала исследования было решено обратиться к устному подкорпусу НКРЯ, очевидным образом также содержащему материалы русской устной речи.

2.3.2. Пользовательский подкорпус НКРЯ

Для того чтобы получить возможность анализировать клаузы, максимально сходные с входящими в состав КРУР, был сформирован пользовательский подкорпус, отвечающий следующим параметрам:

- годы создания: 2005–2017 (т.е. годы создания монологических текстов, входящих в КРУР (<http://russpeech.spbu.ru/source.htm>) и составляющих его большую часть);
- рассматривались только тексты со снятой грамматической омонимией;
- рассматривались только тексты, относящиеся к устной публичной речи.

В результате был получен пользовательский подкорпус НКРЯ объемом 3 документа, 4 957 слов. Материал этого подкорпуса оказался чрезвычайно однородным: так, из выдачи 159 вхождений по запросу *быть* в лексико-грамматическом поиске, 139 примеров приходились на один документ. Для того чтобы не позволить специфике одного конкретного текста влиять на полученную выборку, материал был дополнительно расширен посредством учета текстов, созданных в 2002–2004 гг., причем учитывались первые десять вхождений в каждом из представленных в корпусе документов.

Часть вхождений исключалась из корпуса по тем же параметрам, которые применялись к корпусу русской устной речи. По итогам выборки, пользовательский подкорпус НКРЯ содержал 197 вхождений.

2.3.3. Итоговый рабочий корпус

Полученный итоговый массив рассматриваемых элементарных предложений (далее – просто корпус), полученных из Корпуса устной русской речи и НКРЯ, насчитывает **364** вхождения.

Информация о распределении различных форм глагола *быть* в примерах, вошедших в корпус, приведена в Табл. 4.

Таблица 4. Частотность форм глагола «быть» в материале корпуса

		Форма	Количество вхождений
Инфинитив		быть	32
Форма настоящего времени		есть	121
Формы будущего времени	52	буду	1
		будем	0
		будет	46
		будешь	0
		будете	0
		будут	5
Формы прошедшего времени	158	был	33
		была	32
		было	51
		были	42
Формы императива	1	будь	1
		будьте	0

Как демонстрирует таблица, в итоговом корпусе преобладает подмножество вхождений с формой *есть*; достаточно частотны так же клаузы с формой *будет*, однако другие формы будущего времени или малочисленны, или не встречаются. Формы прошедшего времени стабильно представлены в корпусе, тогда как форма повелительного наклонения встретила только один раз.

Как упоминалось выше, ряд подмножеств вхождений был исключен из объема анализируемых клауз (изначальное количество рассмотренных вхождений – 554), так как функция *быть* в них сведена к чисто грамматической.

Так, вхождения типа (22) очевидным образом должны анализироваться с точки зрения семантических ролей, свойственных актантам и сирконстантам глагола *продолжать*:

(22) (Они никогда ни о чём не договорятся/и **будут продо-
лжать** вас обманывать. [В. Жириновский. Выступление В.
Жириновского в программе «Свобода слова», НТВ (2004) // Интернет] [НКРЯ].

Тем не менее, чрезвычайно высокая частотность таких случаев, а также тот факт, что алгоритм должен быть способен выявить и проанализировать в том числе и клаузы с чисто формальным употреблением *быть*, заставляет нас зафиксировать информацию об их структуре и частотности в рассмотренном материале.

В Табл. 5 приводится информация об абсолютной частотности случаев, входящих в список не рассматриваемых в данной работе подмножеств.

Таблица 5. Частотность исключенных подмножеств

Подмножество	Количество вхождений
Синтетические формы будущего времени изъявительного наклонения	100
Сложные формы страдательного залога	71
Показатели времени/наклонения/модальности при предикате	19

В рассмотренном материале были стабильно представлены следующие случаи формального употребления форм глагола *быть*:

- синтетические формы будущего со следующей структурой:
личная форма глагола быть + инфинитив глагола несовершенного вида: никто с этим не будет спорить [КРУР]
- сложные формы страдательного залога: *личная форма глагола быть + краткое страдательное причастие: после того как наш флот был разрушен в русско-японскую войну* [КРУР]
- показатели времени/наклонения/модальности при предикате:
было + инфинитив глагола: эту школу надо было закрыть [КРУР]

2.4. Модификации инвентаря семантических ролей

2.4.1. Субъект экзистенциального предложения

Входящие в состав обоих подкорпусов клаузы были рассмотрены на предмет реализации семантических ролей, как вошедших, так и не вошедших в опорный список; при этом классифицировались и актанты, и сирконстанты глагола *быть*.

В самом начале работы со сформированными подкорпусами стали очевидны некоторые особенности семантического и синтаксического оформления глагола *быть*, которые должны были неизбежно повлиять на составленную классификацию валентностей.

В первую очередь, эти особенности касаются субъекта (в самом широком смысле) ситуации, описываемой клаузой с глаголом *быть*. Так, в рамках подавляющего большинства существующих классификаций семантических валентностей определяющими параметрами при выделении агенса являются его одушевленность, обладание контролем над ситуацией и осознанное осуществление описываемого действия (см. п. 3).

Однако семантика глагола *быть* также ведет к созданию высказываний, в которых утверждается существование некоего объекта (23), лица (24) или ситуации (25, 26).

(23) *были хорошие школы* [КРУР];

(24) *и в то же время были плохие учителя* [КРУР];

(25) *Нечто похожее было и в Нижегородской области/ где Кириенко с Немцовым имели различные рычаги тогда для влияния на власть* [НКРЯ];

(26) *это вот такой был жизненный урок* [НКРЯ].

Субъект таких экзистенциальных предложений не обладает агентивностью, т. е. не может контролировать описываемую клаузой ситуацию, не инициирует ее начало и часто выражен неодушевленным существительным. Утверждать, что в клаузах такого типа субъект является

пациентом, пассивно претерпевающим свое существование, или стимулом при неупомянута экспериенцере, также представляется необоснованным.

Очевидно, приходится говорить об особом типе субъекта экзистенциального предложения, характерного именно для глагола *быть* (и, возможно, некоторых глаголов становления, ср. *школа стала хорошей*). Этот субъект сильно отличается от определения агенса в клаузах с почти любым другим глаголом, в том числе и при одинаковом синтаксическом оформлении: ср. *он есть* и *он идет*. Предложения с такой семантикой (и, соответственно, наличие их субъекта), по-видимому, универсальны и встречаются во всех языках мира [Lyons 1967].

Тем не менее, нельзя также утверждать, что все клаузы с глаголом *быть* передают экзистенциальную семантику и в них реализуется только описываемый «безагентивный» агенс. Так, в материале корпуса встречаются вхождения, субъект которых вполне возможно отнести к прототипическому агенсу.

Например, в следующих вхождениях *быть* передает значение действия посещения, вполне осознанно совершаемого одушевленными субъектами:

(27) у нас {в} гостях **был** глава *санкт-петербургской общественной организации "семейная коммуна надежда" для подростков и молодёжи, сирот и социальных сирот* [КРУР];

(28) Ребята **были**, действительно, из Японии, из Норвегии, Германии, Бельгии, Польши [КРУР].

Отвечают традиционным параметрам агенса также клаузы, в которых реализуются семантические роли функции (29) или роли (30), (21) (о них см. пункт 2.4.3):

(29) которая раньше была гвардейскими летними лагерями до революции [КРУР];

(30) – я буду прибористом [КРУР];

(31) он никогда не будет... школьной учительницей
[КРУР].

Очевидно, что семантический анализ каждой отдельной клаузы невозможен в рамках поставленной технической задачи, как невозможно и отграничение такого «специфического» агенса от «прототипического» по какому-либо формальному признаку, который может быть учтен алгоритмом. Тем не менее, для того чтобы хотя бы частично разрешить сформулированную выше проблему, а также собрать данные о «прототипическом» и «специфическом» агенсе глагола *быть*, в классификацию был введен распространитель под условным названием *Экзистенциальный Агенс*, под которым далее будет подразумеваться субъект экзистенциального предложения.

Одушевлённые субъекты ситуации, контролирующие ее течение и удовлетворяющие остальным традиционным параметрам Агенса, классифицировались как Агенс.

2.4.2. Объект обладания

Во многом сходную проблему представляет широко распространённое использование глагола *быть* в клаузах с посессивной семантикой, например:

(32) *но есть ли у человечества план "б"?* [КРУР];

(33) *вот, и у нас были совершенно замечательные учителя*
[КРУР].

Как и субъект экзистенциального предложения, субъект предложения, описывающего ситуацию принадлежности, не удовлетворяет параметрам «традиционного» Агенса. Соответственно, для его обозначения была также введена отдельная семантическая роль под условным названием **Объект Обладания**²⁴. Ее выделение, однако, не является таким однозначным ходом,

²⁴ См., опять же, обсуждение этого явления в [Lyons 1967].

как выделение Экзистенциального Субъекта: в то время как границу между Агенсом и Экзистенциальным Субъектом можно провести по ряду очевидных носителю (хотя бы и не компьютерной программе) семантических признаков, разница между Объектом Обладания и Экзистенциальным Субъектом представляется гораздо менее осязаемой.

Рассмотрим, например, следующие вошедшие в корпус клаузы:

(34) *Есть вопросы?* [НКРЯ]

(35) *Если есть вопросы/ [сразу задавайте вопросы]*
[НКРЯ]

(36) *И тут было мнение/ [что НАТО/ это хорошо]*
[НКРЯ]

(37) *Когда бы был юрист и были бы деньги, [мы бы]...*
[НКРЯ]

Наивному носителю языка достаточно очевидно, что в ситуациях, описанных клаузами (34), (35), у выделенных Объектов Обладания есть невыраженный посессор, являющийся адресатом высказывания; в высказываниях (36), (37) также подразумевается наличие посессора: так, мнение (особенно в рамках ограниченной по времени и месту ситуации (*тут*)) не может никому не принадлежать, общий контекст высказывания (37) также подразумевает наличие некоего коллективного посессора, с которым отождествляет себя говорящий. Ср. след. примеры, где в клаузы 24–37 добавлены эксплицитные выражения посессоров:

(38) ***У вас** есть вопросы?*

(39) *Если **у вас** есть вопросы/ [сразу задавайте вопросы]*

(40) *И тут **у** **кого-**
то было мнение/ [что НАТО/ это хорошо]*

(41) *Когда бы **у нас** был юрист и были бы деньги, [мы бы]...*
[КРУР]

Очевидно, что во многих контекстах автоматическое отделение Объекта Обладания от Экзистенциального Субъекта (как семантических сущностей) невозможно. Для разрешения этой проблемы был введен формальный критерий наличия **реализованного Посессора**. Т. е. в контекстах типа (34–35) обсуждаемый участник ситуации трактовался как Экзистенциальный Субъект в тех случаях, когда в той же клаузе не было актанта или сирконстанта, заполняющего роль Посессора. В тех же случаях, где Посессор был эксплицитно реализован, обсуждаемый участник ситуации классифицировался как **Объект Обладания**. См., например, следующие вхождения, где семантическая роль Посессора реализуется посредством ввода отдельного распространителя:

(42) *...у тебя есть что-то добавить?* [НКРЯ]

(43) *И для того чтобы у нас было нормальное понимание всего этого...* [НКРЯ].

При желании, из этих клауз может быть устранён Посессор, что не повлечёт за собой радикальных изменений в значении клаузы или плана значения и выражения Объекта Обладания. Ср. следующие модифицированные примеры:

(44) *... есть что-то добавить?*

(45) *И для того, чтобы было нормальное понимание всего этого...*

Таким образом, в контексте прикладных задач данного исследования представляется важным выделить дополнительную семантическую роль Объекта Обладания. Однако, следует иметь ввиду, что ее отделение от семантической роли Экзистенциального Субъекта в контексте данной работы является во многом условным.

2.4.3. Определение, Определение: как, Состояние, Роль и Функция

В п. 3 описывалась выделяемая в [Апресян и др. 2014] функция быть 1.1 в значении 'являться', которая объединяет такие случаи употребления, как:

(46) *Мой сын будет архитектором*

(47) *Парень был слегка навеселе*

(48) *Он в восторге от твоего доклада* [Апресян и др. 2014: 394].

Однако представляется, что эта группа требует дополнительной детализации с т.з. семантики «состояния или свойства» [Там же], которым характеризуется описываемый участник. Едва ли решение (как правило, добровольное) о том, чтобы выбрать некую профессию (46), может быть отождествлено с ситуацией ощущения некой эмоции (47): например, достаточно общепринято разделение на Агенс и Экспериенцер, в терминах которых описывают субъекты первой и второй ситуаций соответственно.

Для разграничения участников, описывающих уже сами свойства субъекта, были введены дополнительные семантические роли со следующими обобщенными значениями:

- **Состояние:** см. п. 1.5.2.

- **Определение:** характеристика лица объекта, являющегося участником ситуации, и описывающего его с точки зрения всех свойств, не охваченных семантической ролью Состояние.

(49) *покой был **известного рода*** [КРУР];

(50) *они должны пользоваться поддержкой государства или быть **государственными*** [КРУР];

- **Определение: как:** характеристика описываемой клаузой ситуации в целом. Как правило, используется в безличных клаузах:

(51) *и это было **так ужасно*** [КРУР];

(52) *Так что/ Сергей/ я уверена/ всё будет **нормально*** [НКРЯ].

Очевидно, что грань между семантическими ролями Определение: как? и Состояние может быть существенно размытой. Так, например, ср. следующие два вхождения:

(53) *потому что были не раз случаи/ когда я была рядом с ней/ когда ей было **тяжело**/* [НКРЯ];

(54) *уже очень **тяжело** будет* [НКРЯ].

Интуитивно понятно, что в ситуации, описываемой второй клаузой, также существует тот или те, *для кого* ситуация будет тяжелой; таким образом, на уровне семантики характеристика ситуации часто взаимосвязана (или даже взаимозаменяема) с состоянием ее одушевленных участников.

Для в будущем автоматического разграничения этих двух семантических ролей был введен формальный критерий эксплицитной выраженности Экспериенцера, подразумевающий, что при наличии распространителя с такой ролью случай трактовался как Состояние.

▪ **Роль:** характеристика одушевленного объекта ситуации, описывающая его с т.з. «работы, рода деятельности, проявления себя в качестве кого-л» [МАС]; таким образом, обладание свойством, описываемым с помощью Роли, обычно ограничено во времени:

(55) *мечтал стать юным **ленинцем** и быть членом лучшей октябрятской звёздочки* [КРУР];

(56) *Нужно быть **фаталистом**/ недаром мы в японском р есторане сидим* [НКРЯ].

▪ **Функция:** характеристика неодушевленного объекта ситуации, описывающая его с т.з. «работы, рода деятельности, проявления себя в качестве кого-л». Семантическая роль Функции отделялась от Роли только по признаку одушевленности субъекта клаузы и, соответственно, часто (но не всегда) его волитивности;

(57) *которая раньше была **гвардейскими летними лагерями** до революции* [КРУР];

(58) *что это будет **догмой*** [НКРЯ].

2.4.4. Тожественность

Еще одна функция быть, выделяемая в [Апресян и др. 2014], – это *быть* 1.2 в значении 'отождествление': «A1 есть A2» [Апресян и др. 2014: 395]. Для передачи этой функции в контексте нашего инвентаря была введена семантическая роль **Тожественности**, которая кодировала участника A2, тогда участник A1 классифицировался как Экзистенциальный Агенс. См. следующие вхождения:

(59) *это был наш основной **принцип*** [КРУР];

(60) *это была **другая ситуация*** [КРУР];

(61) *воспитание есть **процесс обучения**, приведения к балансу социального статуса и уровня социальных притязаний* [КРУР].

Итак, на этом этапе работы опорный инвентарь семантических ролей был модифицирован таким образом, чтобы удовлетворять нуждам описания функционирования глагола *быть*. Результатам описания посвящен следующий пункт работы.

2.5. Частотность активации семантических ролей

2.5.1. Общая информация

Общее количество распространителей глагола *быть*, заполнявших его семантические валентности в проанализированном материале, составило **678** единиц. Относительная частотность реализации каждой семантической роли отсчитывалась от этого числа. Всего в проанализированном материале были выявлены **32** реализованные семантические роли. Рисунок 1 демонстрирует абсолютную частотность случаев их активации в клаузах с глаголом *быть*.

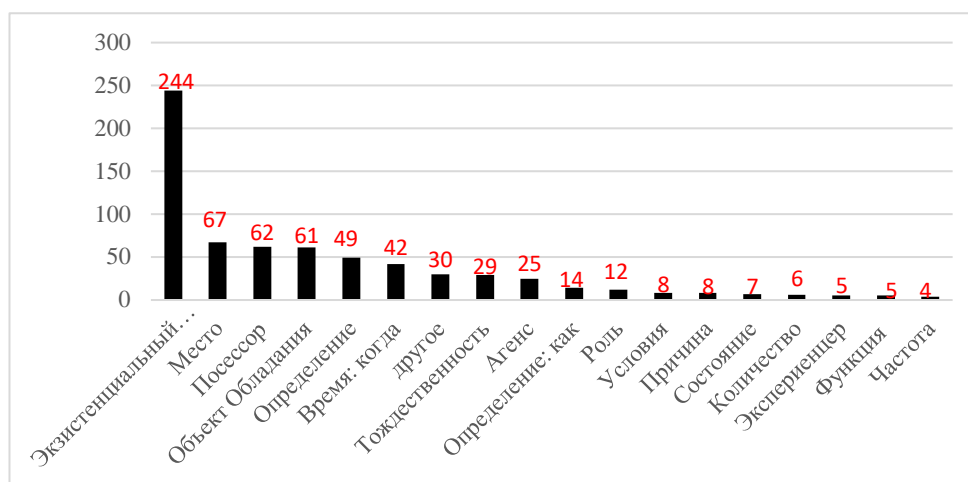


Рисунок 1. Частота реализации семантических ролей

Итак, наиболее частотно активирующейся семантической ролью является Экзистенциальный субъект; за ним следуют Место, Посессор, Объект Обладания, Определение и Время: когда.

«Прототипический» Агнс встречается не так часто. Здесь же следует упомянуть, что Пациенс отсутствует во всем проанализированном материале. Важно, однако, отметить, что последний факт очевидным образом объясняется исключением из рассматриваемого массива случаев употребления форм глагола *быть* в сложных формах страдательного залога, в целом достаточно частотных в русской устной речи (см. п. 2. 3.1).

В категорию «Другое» были отнесены семантические роли, которые активировались только в одной–трех клаузах. Информация о единично представленных семантических ролях приведена на Рисунке 2.

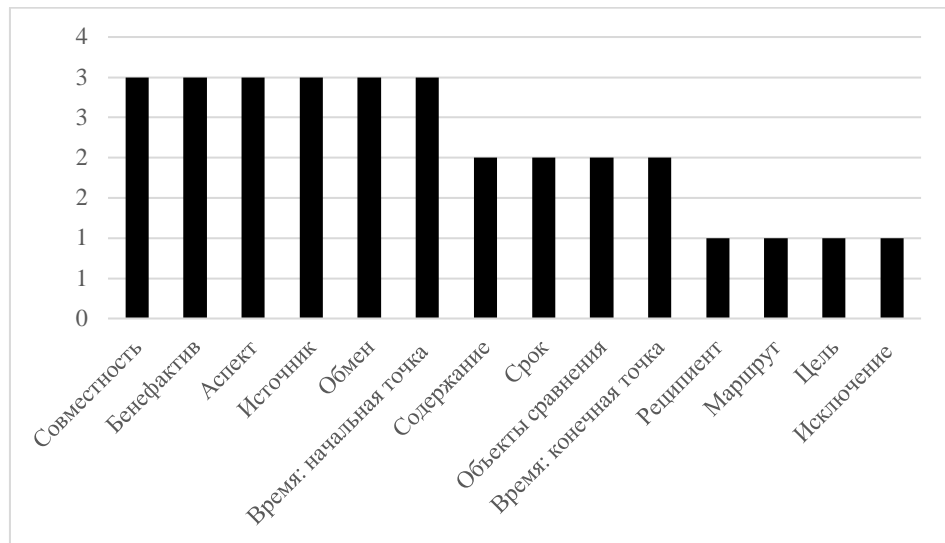


Рисунок 2. Частотность реализации семантических ролей: другое

Итак, случаи активации различных семантических ролей представлены в корпусе крайне неравномерно; в целом, уже на этом этапе можно говорить о наличии высоко- и нетипичных для глагола *быть* ролей распространителей.

2.5.2. Сочетаемость семантических ролей

На следующем этапе исследования анализировалась сочетаемость семантических ролей. Здесь и далее под *сочетаемостью* будет подразумеваться способность нескольких семантических ролей одновременно реализовываться в пределах одной клаузы.

В корпусном материале максимальное число распространителей, выполняющих разные семантические роли в одной клаузе, составило 4, минимальное – 0.

В п. 3.2.1-3.2.4 последовательно обсуждаются такие случаи в порядке от большего к меньшему. Не рассматривались вхождения, в которых были эксплицитные указания на незавершенность клаузы (например, многоточие в расшифровке):

(62) она была...

(63) и были такие замечательные у нас...

(64) Я не знаю/ что и ответить... раньше был... ммм...

Представляется, что для полного описания таких случаев необходим также анализ их интонационного контура.

Также исключались случаи с высоко специфичным значением, оказывающим влияние на грамматическую структуру предложения: например, указание на возраст какого-либо участника ситуации.

(65) *мне было семнадцать или восемнадцать лет.*

Реализация четырех семантических ролей в пределах одной клаузы

В проанализированном материале корпуса встречаются всего пять вхождений, в которых заполняются одновременно четыре валентности глагола *быть*. Четыре клаузы имеют общую семантическую структуру **Объект Обладания + Посессор + Время: когда + Семантическая роль X²⁵**.

В первом таком вхождении слот X заполняется семантической ролью Количества (66), во втором – семантической ролью Условий (67).

(66) *у нас (Посессор) других счастливых воспоминаний (Объект Обладания), на самом деле, не так много (Количество) было в советское время (Время: когда) [КРУР]*

(67) *сейчас (Время: когда) у санкт-петербургских школьников (Посессор) также есть возможность почувствовать вкус технического творчества (Объект Обладания) при работе над собственными космическими проектами (Условия) [КРУР]*

²⁵ Здесь и далее список семантических ролей приводится без учета порядка следования выполняющих их распространителей, соответственно условны способы референции к ним типа «вторая семантическая роль».

Два других случая, в которых распространители выполняют семантическую роль Места, требуют более детального рассмотрения.

(68) *который (Экзистенциальный Субъект) был у нас (Посессор) {в} стране (Место) в шестидесятые годы (Время: когда)*
[КРУР]

(69) *[Представляю]/ что сейчас (Время: когда) будет у нас (Посессор) в зале (Место) / номер одиннадцатый/ Дмитрий Сычев (Объект обладания)!* [НКРЯ]

Так, комбинация, обсуждаемая здесь и далее в терминах сочетания Места и Посессора²⁶, часто может быть истолкована и как один распространенный актант, где указание на принадлежность места кому-либо является уточнением значения места, а не отдельной семантической единицей. Однако вполне допустимыми (особенно в устной речи) представляются и модифицированные варианты этих клауз:

(70) *который (Экзистенциальный Субъект) был у нас (Посессор) в шестидесятые годы (Время: когда)* [КРУР]

(71) *[Представляю]/ что сейчас у нас (Посессор) будет / номер одиннадцатый/ Дмитрий Сычев (Объект обладания)!*

Ср., например, со следующим вхождением из [КРУР] (72) и его вариантом, модифицированным «в обратную сторону», т.е. посредством добавления дополнительного распространителя, выполняющего семантическую роль Места (73):

(72) *и у нас (Посессор) были замечательные учителя (Объект Обладания)* [КРУР]

(73) *и у нас (Посессор) в школе [Место] были замечательные учителя (Объект Обладания)* [КРУР]

Очевидно, что неоднозначность в интерпретации семантической структуры таких вхождений вызвана тем, что фактический посессор и Места

²⁶ См. также обсуждение более редкого сочетания Посессор + Условия п. 3.2.2.

(например, обозначаемого распространителями *страна* в 68, *зал* в 69, *школа* в 73), и Объекта обладания (соответственно, *которые* в 68, 70, *Номер одиннадцатый, Дмитрий Сычев* в 69, 71, *замечательные учителя* в 72, 73) совпадает (во всех перечисленных случаях это коллективный говорящий, к которому реферирует распространитель *у нас*), и, соответственно, точно определить семантические отношения элементов клаузы не представляется возможным.

Ср., например, с вхождением, где Посессор не может относиться к месту, т. к. Казань определенно не принадлежит адресату высказывания:

(74) *а у вас есть в Казани друзья?* [НКРЯ].

Итак, определить, к кому или чему именно происходит референция во вхождениях типа 66–67 и, соответственно, точно их интерпретировать не представлялось возможным. При этом, как будет показано далее, они очень частотны в проанализированном материале корпуса и, следовательно, должны быть зафиксированы в описании, на которое будет опираться алгоритм. Итак, было принято решение о выделении в описанных неоднозначных случаях двух семантических ролей Места и Посессора. Тем не менее, с точки зрения чисто семантической интерпретации клауз, следует помнить, что в некоторых вхождениях это разделение может быть избыточным.

Еще одно вхождение, в котором реализуются одновременно четыре семантические роли, имеет структуру **Экзистенциальный субъект + Функция + Время: когда + Время: конечная точка**.

(75) *которая (Экзистенциальный субъект) раньше (Время: когда) была гвардейскими летними лагерями (Функция) до революции (Время: конечная точка)* [КРУР]

Реализация трех семантических ролей в пределах одной клаузы

Случаи активации трех семантических ролей в пределах одной клаузы гораздо более частотны и наблюдаются в 45 вхождениях. В этом множестве можно выделить три крупные подгруппы, семантическая структура клауз в которых может быть описана следующим образом:

- **Экзистенциальный субъект + Семантическая роль X + Семантическая роль Y** (21 вхождение);
- **Объект Обладания + Посессор + Семантическая роль X** (18 вхождений);
- **Агнс + Семантическая роль X + Семантическая роль Y** (4 вхождения).

Далее последовательно рассмотрим эти подгруппы, а также не вошедшие в них клаузы.

▪ **Экзистенциальный субъект + Семантическая роль X + Семантическая роль Y**

В подмножестве клауз с этой семантической структурой можно выделить крупную подгруппу (9 вхождений), в которой слот второй семантической роли X занимает Определение. В трех таких случаях семантической ролью Y является Время: когда (76), в трех – Условия (77), в двух – Место (78).

(76) *То есть вы (Экзистенциальный субъект) раньше (Время: когда) были такой медитативный и какой-то светлый (Определение) [НКРЯ]*

(77) *То тогда (Условия) эта цифра (Экзистенциальный субъект) будет ещё меньше (Определение) [НКРЯ]*

(78) *А / например / в Псковской области (Место) ситуация (Экзистенциальный субъект) была ещё интереснее (Определение) [НКРЯ]*

В одном вхождении третьей реализующейся семантической ролью

оказывается **Время: начальная точка**.

(79) *После этого (Время: начальная точка) примерно рост (Экзистенциальный субъект) был символический (Определение)*
[НКРЯ]

Структуру следующей по объему подгруппы (4 вхождения) можно описать как **Экзистенциальный Субъект + Тождественность + Семантическая роль Y**. В трех таких случаях слот Y заполняет распространитель, выполняющий семантическую роль **Время: когда** (80), в одном – **Обмена** (81).

(80) *Это (Экзистенциальный субъект) был разговор (Тождественность) в декабре месяце (Время: когда)* [НКРЯ]

(81) *будь это (Экзистенциальный субъект) занятия (Тождественность) за деньги (Обмен)...* [КРУР]

Четыре вхождения приходится на комбинацию **Экзистенциальный Субъект + Время: когда + Семантическая роль Y**, где в двух случаях слот Y заполняет **Место**, а в двух других - **Маршрут** и **Срок**.

(82) *недавно (Время: когда) была встреча (Экзистенциальный субъект) в администрации кировского района (Место)* [КРУР]

(83) *переноска грузов (Экзистенциальный субъект) будет уже (Время: когда) не по рельефу/ а по навешенному/ обработанному пути/ (Маршрут)* [НКРЯ]

(84) *а потом (Время: когда) было "окно" (Экзистенциальный субъект) дней пять-шесть (Срок)* [НКРЯ]

Вне перечисленных подгрупп остается ряд клауз, в которых представлены единичные случаи сочетаемости семантических ролей.

Так, структура **Экзистенциальный Субъект + Место + Количество**

(85) встречается два раза, структуры **Экзистенциальный Субъект + Реципиент + Обмен** (86) и **Экзистенциальный Субъект + Бенефактив + Аспект** (87) – по одному разу.

(85) *сколько* (Количество) *их* (Экзистенциальный субъект) *там* (Место) *есть* [КРУР]

(86) [*Думаешь*]/ *тебе* (Реципиент) *мороженое* (Экзистенциальный субъект) *за первое место* (Обмен) *будет?* [КРУР]

(87) *Чтобы* *ко мне* (Бенефактив) *было доверие* (Экзистенциальный субъект) *как к врачу* (Аспект) [НКРЯ]

▪ **Объект Владения + Посессор + Семантическая роль X**

В пяти вхождениях слот X заполняется семантической ролью Места, причем среди них есть случаи как содержащие (88), так и не содержащие описанную выше неоднозначность (89).

(88) *А есть ли вообще у нас* (Посессор) *в правительстве* (Место) *такая фигура* (Объект владения)

(89) *Это как у нас* (Посессор) *на Эвересте* (Место) *было сразу восемь экспедиций* (Объект владения) [НКРЯ]

По два вхождения приходилось на каждый из следующих вариантов заполнения слота X: семантическая роль Определение (90) и Причина (91).

(90) *Но обувь и личное снаряжение* (Объект Владения) *у вас* (Посессор) *должны быть нормальными* (Определение) [НКРЯ]

(91) *Но поскольку у нас* (Посессор) *на самом деле по поводу каравана* (Причина) *есть сомнения и большие проблемы* (Объект владения) [НКРЯ]

Единично представлены вхождения, где третий распространитель выполняет семантическую роль Условие, Время: когда, Время: конечная точка, Время: начальная точка, Срок, Совместность, Аспект, Содержание и Цель. См., например:

(92) *У нас* (Посессор) *был определённый рывок* (Объект

обладания) до 94-го года (Время: конечная точка) [НКРЯ]

(93) Но и у многих из вас (Посессор) есть слабости (Объект обладания) не в одном/ так в другом (Аспект) [НКРЯ]

(94) После переправы через реки (Время: начальная точка) у нас (Посессор) никакой связи с внешним миром (Объект Обладания) не будет [НКРЯ]

▪ **Агенс + Семантическая роль X + Семантическая роль Y**

Три «трехместных» вхождения с семантической структурой Агенс + Семантическая роль X + Семантическая роль Y объединяет и второй элемент: распространитель, выполняющий семантическую роль Места.

В двух случаях слот Y заполняется в одном случае семантической ролью Время: когда, во третьем – семантической ролью Частота:

(95) Ни разу (Частота) не были они (Агенс) в Нижнем Новгороде (Место)... [НКРЯ]

(96) Когда недавно (Время: когда) Олег Медведев (Агенс) был в Казани (Место) [НКРЯ]

(97) Я (Агенс) 13-го (Время: когда) должен быть на работе (Место) [НКРЯ]

Третьим в этой группе был во многом нетривиальный пример единично встретившейся в материале корпуса семантической структуры, которая была интерпретирована как комбинация семантических ролей **Агенс + Посессор + Условия:**

(98) у нас (Посессор) {в} гостях (Условия) был глава *санкт-петербургской общественной организации "семейная коммуна надежда"* для подростков и молодёжи, сирот и социальных сирот (Агенс) [КРУР]

Во-первых, здесь мы снова сталкиваемся со сложностями, связанными с однозначным установлением объекта референции Посессора (см. п. 3.2.1), однако в этот раз встает вопрос о разделении Посессора и Условий, а не

Посессора и Места.

Кроме того, вызывает вопросы и сама возможность говорить о принадлежности волитивного и обладающего контролем Агенса какому-либо другому участнику ситуации, пусть даже и в узко специфическом контексте. Например, возможна также трактовка, в рамках которой распространитель у нас выполняет семантическую роль Места, указывая на локализацию описываемой ситуации в пространстве (например, если говорящий подразумевает офис радиостанции, который посетил Агенс). Однако предложенный в п. 3.2.1 тест показывает возможность присоединения дополнительного сирконстанта, семантическая роль которого однозначно интерпретируется как Место:

(99) у нас (Посессор) на радиостанции (Место) {в} гостях (Условия) был глава санкт-петербургской общественной организации "семейная коммуна надежда" для подростков и молодёжи, сирот и социальных сирот (Агенс)

Таким образом, представляется возможным говорить о сочетаемости Агенса и Посессора, по крайней мере во клаузах, где формы глагола *быть* имеют значение быть 2.1, быть 2.2 [Апресян и др. 2014] или быть 2, 3 [МАС] (см. п. 1.6).

▪ Другое

Два других вхождения, где активируются 3 семантических роли, имеют семантическую структуру **Экспериенцер + Состояние + Причина** и **Определение: как + Место + Время: когда**.

(100) Поэтому (Причина) нам (Экспериенцер) будет очень непросто (Состояние) [НКРЯ]

(101) поскольку там (Место) было всё время (Время: когда) плохо/ плохо/ плохо/ (Определение: как) [НКРЯ]

Реализация двух семантических ролей в пределах одной клаузы

Самыми частотными в проанализированном материале корпуса были случаи одновременной реализации в пределах одной клаузы двух семантических ролей. Общее количество таких вхождений составило **212**.

Внутри этого подмножества с точки зрения семантической структуры можно выделить семь крупных подгрупп частотных клауз, в пределах каждой из которых сочетается одна пара семантических ролей: Экзистенциальный Субъект + Место, Объект обладания + Посессор, Экзистенциальный Субъект + Определение, Экзистенциальный Субъект + Тождественность, Экзистенциальный Субъект + Время: когда, Агенса + Роль, Агенса + Место.

Далее приведем примеры таких наиболее частотных «двухместных» комбинаций:

– **Экзистенциальный Субъект + Место** (42 вхождения)

(102) *то есть там (Место) была подготовка*
(Экзистенциальный субъект) [КРУР]

(103) *в Норвегии (Место) тоже есть космодром*
(Экзистенциальный субъект) [КРУР]

– **Объект обладания + Посессор** (36 вхождений)

(104) *вот у нас (Посессор) была ситуация (Объект
обладания)* [КРУР]

(105) *У Комиссарова (Посессор) есть маленькая*
фотография (Объект обладания) [НКРЯ]

– **Экзистенциальный Субъект + Определение** (30 вхождений)

(106) *такова (Определение) была задача той школы*
(Экзистенциальный субъект) [КРУР]

(107) *Тебе (Экзистенциальный субъект) надо быть*

неприступной (Определение). [НКРЯ]

– **Экзистенциальный Субъект + Тожественность** (28 вхождений)

(108) но она (Экзистенциальный субъект) была не пример для подражания (Тожественность) [КРУР]

(109) [Для того чтобы вам было понятно]/ кто (Экзистенциальный субъект) есть кто (Тожественность) [НКРЯ]

– **Экзистенциальный Субъект + Время: когда** (15 вхождений)

(110) что в советские времена (Время: когда) была прекрасная школа (Экзистенциальный субъект) [КРУР]

(111) Когда (Время: когда) никакой всероссийской славы (Экзистенциальный субъект) не было. [НКРЯ]

– **Агенси + Роль** (9 вхождений)

(112) хотя моя бабушка (Агенси) не была школьной учительницей (Роль). [КРУР]

(113) каждый первоклассник (Агенси) [мечтал стать юным ленинцем]и быть членом лучшей октябрятской звёздочки (Роль). [КРУР]

– **Агенси + Место** (6 вхождений)

(114) Которые (Агенси) были в районе Эвереста (Место) [и представляют всё это] [НКРЯ]

(115) и мы (Агенси) хотим быть в русле современной жизни (Место) [КРУР]

Часть клауз, не вошедших в перечисленные группы, можно классифицировать следующим образом с точки зрения их семантической структуры:

- **Экзистенциальный субъект + Семантическая Роль X** (27 вхождений), где слот X заполняется 14 разными семантическими ролями. Сведения о способах его заполнения X и их частотности приведены на рис. 3.

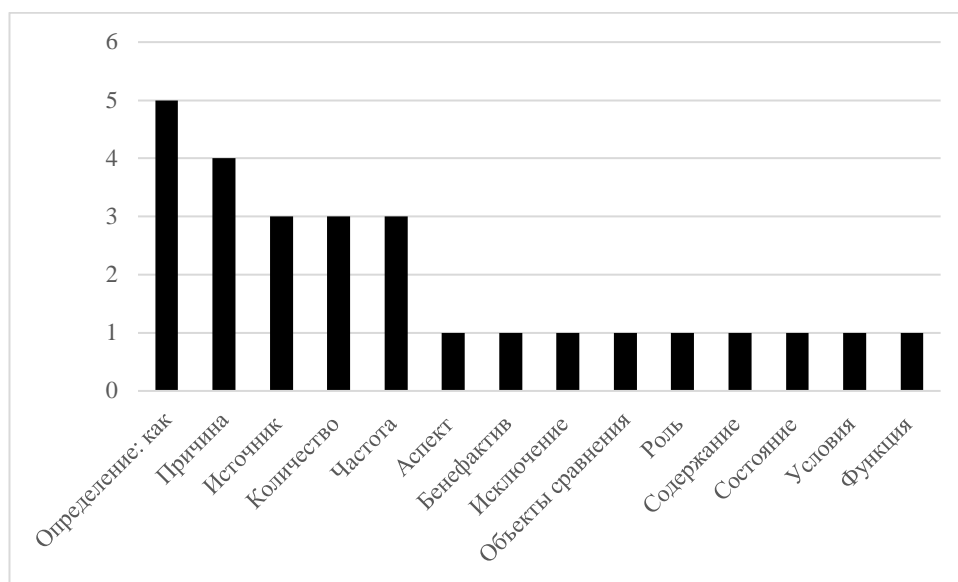


Рисунок 3. Частотность заполнения слота X

См., например, следующие вхождения:

(116) *всё* (Экзистенциальный субъект) *будет нормально*
(Определение: как) [НКРЯ]

(117) *Это* (Экзистенциальный Субъект) *будет/ конечно/ более эффективно* (Определение: как) [НКРЯ]

(113) *Поэтому* (Причина) *есть такое предложение*
(Экзистенциальный субъект) [НКРЯ]

- **Агентс + Семантическая роль X** (5 вхождений), где слот X заполняется 4 разными семантическими ролями.

Чаще всего в таких клаузах второй заполненной ролью оказывалась семантическая роль Определение (2 вхождения); остальные опции (Время: когда, Совместность, Условия) встречались единично.

(118) *Ребята* (Агентс) *были, действительно, из Японии, из*

Норвегии, Германии, Бельгии, Польши (Определение). [КРУР]

(119) *и Глазьев* (Агенс) *был с Ельциным* (Совместность)
[НКРЯ]

Кроме того, в корпусном материале встречается ряд двухместных клауз, не вошедших ни в одно из перечисленных подмножеств. С точки зрения семантической структуры их можно описать как **Определение**: как **Время**: когда, **Бенефактив** + **Функция**, **Место** + **Время**: когда, **Объект** сравнения + **Определение**, **Посессор** + **Количество**, **Совместность** + **место**, **Частота** + **Условия**, **Экспериенцер** + **Состояние**, **Определение** + **Обмен**.

Все случаи сочетаемости перечисленных пар семантических ролей представлены в корпусном материале единично; исключение составляет комбинация **Определение**: как + **Время**: когда, возникающая в трех вхождениях. См., например:

(120) *всегда* (Время: когда) *так* (Определение: как) *было*.
[КРУР]

Реализация одной семантической роли в пределах одной клаузы

Количество вхождений, содержавших только один распространитель, выполнявший, соответственно, одну семантическую роль, составило **88**.

В абсолютном большинстве таких случаев (81 вхождение) была реализована семантическая роль **Экзистенциального субъекта**.

(121) *были хорошие школы* (Экзистенциальный субъект)
[КРУР]

(122) *есть руководство экспедиции* (Экзистенциальный субъект) [НКРЯ]

(123) *есть руководитель* (Экзистенциальный субъект)
[НКРЯ]

Остальные семь случаев составили клаузы, где реализовывались следующие семантические роли: Определение и Определение: как (по 2 вхождению), Посессор, Роль, Время: когда (по 1 вхождению).

Реализация 0 семантических ролей в пределах одной клаузы

Кроме исключенных случаев с эксплицитным выражением незавершенности клаузы, в материале корпуса наблюдались и вполне завершенные с т. з. интонационного контура высказывания, в которых не реализовывалось ни одной семантической роли.

Это подмножество составляли ответные реплики, в которых содержались отсылки к предыдущим клаузам в общем массиве диалога:

(121) *Конечно/ есть.* [НКРЯ]

(122) *И не было.* [НКРЯ]

Таким образом, представляется, что в таких случаях можно говорить о контекстно обусловленном опущении распространителей, заполняющих синтаксические валентности глагола *быть* при сохранении его семантических валентностей.

Краткое обобщение

В этом параграфе обсуждались случаи сочетаемости семантических ролей глагола *быть*. С точки зрения составления опорной базы данных для функционирования алгоритма восприятия речи эта информация должна быть представлена именно в формате количества семантических ролей, реализующихся в пределах одной клаузы.

Тем не менее, нельзя не отметить некоторые общие тенденции в области сочетаемости семантических ролей, наблюдаемые при рассмотрении материала корпуса в целом. Так, наиболее частотными оказывались клаузы, в пределах которых реализовывались одновременно две семантических роли. Следующими по частотности были вхождения с активацией только одной роли; наиболее редко встречались случаи, где активировалось 4 или 0 ролей.

Кроме того, представляется важным зафиксировать, что следующие комбинации семантических ролей наиболее склонны активироваться в пределах одной клаузы²⁷:

- Объект Обладания + Посессор;
- Экзистенциальный субъект + Определение;
- Экзистенциальный субъект + Место;
- Экзистенциальный субъект;
- Экзистенциальный субъект + Тождественность;
- Агенс + Роль;
- Агенс + Место;
- Место + Посессор.

2.6. Распределение семантических ролей по формам глагола *быть*

Выше обсуждалась сочетаемость семантических ролей между собой в пределах одной клаузы. Однако в контексте функционирования алгоритма, в котором предусмотрено, что в первую очередь анализируется предикат клаузы, и только потом – его зависимые, большое значение имеет и сочетаемость семантических ролей, активирующихся в клаузах, с формами глагола *быть*, а также описание того, чем именно оказываются выражены семантические роли.

Учитывая, что различные формы глагола *быть*, так же, как и распространители, выполняющие различные семантические роли, представлены в материале корпуса чрезвычайно неравномерно, представляется необходимым привести сведения о сочетаемости между ними со следующих двух ракурсов:

- Частотность форм глагола *быть* в клаузах с активированной частотной семантической ролью X;
- Частотные случаи активации семантических ролей в клаузах с Y формой глагола *быть*.

²⁷ Приводятся не в порядке частотности.

Для представления данных первого рода были отобраны семантические роли, которые были реализованы в материале более 15 раз. Согласно п. 3.1, в этот список вошли следующие роли:

- Экзистенциальный субъект
- Место
- Посессор
- Объект Обладания
- Определение
- Время: когда
- Тождественность
- Агенс

Информация о распределении распространителей с перечисленным семантическими ролями по формам глагола *быть* представлена в Приложениях 3–10; частотность семантических ролей в клаузах с формами инфинитива, прошедшего и будущего времени *быть*, а также формы *есть*, однако, заслуживает отдельного комментария.

2.6.1. Типичные случаи активации семантических ролей в клаузах с «*быть*»

В клаузах с формой *быть* наиболее часто реализуются семантические роли Экзистенциального субъекта (18 вхождений) и Определения (13 вхождений)²⁸.

Семантическая роль **Экзистенциального субъекта** может быть заполнена именной группой, а также единичным существительным, прилагательным или указательным местоимением²⁹; при этом падеж распространителя зависит от модального глагола, которому подчинен инфинитив. См., например:

²⁸ Здесь и далее такие общие по форме утверждения относятся только к материалу корпуса.

²⁹ Далее случаи употребления единичных и распространенных существительных, прилагательных и местоимений будут объединяться под общим понятием Именной группы.

(124) *Тебе надо быть неприступной.* [НКРЯ]

(125) *Здесь **никаких послаблений** ни себе/ ни своим близка
йшим друзьям быть не может.* [НКРЯ]

Распространитель, выполняющую семантическую роль **Определения**, может быть выражен полным или кратким прилагательным (или именной группой с полным или кратким прилагательным) в Творительном или Именительном падеже:

(126) *правила записи языковых единиц с помощью фонетической, фонологической или морфонологической транскрипции теоретически могут быть **независимы от орфографии*** [КРУР]

(127) *а **какие** могут быть последствия такой амнистии*
[НКРЯ]

2.6.2. Типичные случаи активации семантических ролей в клаузах с «*есть*»

Наиболее частотно сочетающейся с формой *есть* семантической ролью оказывается, опять же, **Экзистенциальный субъект**: в корпусе насчитывается 90 вхождений с такой комбинацией. Однако в случае клауз с этой глагольной формой прослеживается прямая зависимость между ней и способом выражения распространителя, выполняющего семантическую роль Экзистенциального субъекта. Так, во всех случаях, когда на рассматриваемую именную группу не влияет конструкция с числительным (т. е. кроме 3-х вхождений³⁰), распространитель (или главное слово в нем) имеет форму Именительного падежа:

(128) *Я хочу сказать/ что в предстоящем сезоне так же есть ещё два резервных варианта* [НКРЯ]

(129) *Есть новые условия* [КРУР]

Отдельно следует отметить подгруппу из 6 вхождений с указательным

³⁰ См. также Приложение 2.

местоимением *это* в Именительном падеже, которая представляет собой довольно редкий пример реализации семантической роли посредством одной и той же словоформы.

Следующую по объему группу семантических ролей, частотно сочетающихся с *есть*, составляет комбинация **Объект Обладания + Посессор**, возникающая в 25 вхождениях. При этом Объект Обладания всегда кодируется Именительным падежом, а Посессор – именной группой с существительным, личным местоимением или прилагательным в родительном падеже, имеющей структуру *у + ИГ*:

(130) *просто у нас (Посессор) есть один мальчик (Объект Обладания)* [КРУР]

(131) *У кого (Посессор) есть возможность и силы (Объект Обладания)* [КРУР]

В материале корпуса встретилось только одно вхождение, в котором Посессор был выражен именной группой в Родительном падеже, но без предлога *у*:

(132) *есть ещё юридическая сторона (Объект обладания) этого дела (Посессор)* [НКРЯ]

Третьей по частотности активации в клаузах с *есть* семантической ролью является **Тождественность** (как было показано выше, всегда сочетающаяся с Экзистенциальным Субъектом). Она представлена в 10 вхождениях корпуса и всегда кодируется именной группой в Именительном падеже:

(133) *воспитание (Экзистенциальный субъект) есть формирование ценностей, жизненных установок человека (Тождество).* [КРУР]

2.6.3. Типичные случаи активации семантических ролей в клаузах с формами будущего времени «быть»

Семантической ролью, наиболее частотно реализующейся в клаузах

будущего времени, ожидаемо оказывается, **Экзистенциальный субъект** (являющийся, как было показано выше, самой частотной семантической ролью в корпусе). Он активируется в 35 вхождениях, в большинстве (из которых (25) соответствующий распространитель представляет собой именную группу, кодируемую Именительным падежом. Однако в клаузах с негативной формой глагола *быть*, а также в случаях, когда именная группа включает в себя числительное, оказывающее влияние на падеж, распространитель может принимать и форму Родительного падежа (10 вхождений).

(134) *Сегодня будет **много** всего интересного*

(Экзистенциальный субъект) [НКРЯ]

(135) *если не будет педагога, воспитателя*

(Экзистенциальный субъект). [КРУР]

Следующее по объему (8 вхождений) подмножество образует комбинация **Объект Обладания + Посессор**, где Объект обладания кодируется так же, как и Экзистенциальный субъект при форме *есть*, а Посессор – так же, как в п. 4.2.

(136) *что внутренняя связь (Объект _____ обладания)*

у нас (Посессор) будет. [НКРЯ]

2.6.4. Типичные случаи активации семантических ролей в клаузах с формами прошедшего времени «быть»

Подгруппа клауз с формами прошедшего времени глагола *быть* составляет, как было показано в п. 2.3.3, большую часть корпуса и, соответственно, во многом определяет общие тенденции сочетаемости форм глагола *быть* и семантических ролей, представленные в Приложениях 3-10, а также тенденции сочетаемости между семантическими ролями, описанные в п. 3.

Ожидаемо, наиболее частотно сочетающейся с формами прошедшего времени семантической ролью является Экзистенциальный субъект (78 вхождений); комбинации **Экзистенциальный субъект + Определение** и **Экзистенциальный субъект + Тождественность** встречаются в 24 и 17 вхождениях соответственно и кодируются по описанным в п. 4.1-4.2 закономерностям.

2.7. Способы кодирования наиболее абсолютно частотных семантических ролей в клаузах с глаголом *быть*

Выше обсуждались наиболее типичные случаи сочетаемости форм глагола *быть* с семантическими ролями, а также способы выражения последних. Однако, как было показано, вследствие «предпочтения» некоторых семантических ролей и их комбинаций формами глагола *быть*, несколько достаточно частотно реализующихся семантических ролей не были описаны.

Итак, в этом пункте будут кратко освещены основные способы кодирования частотных семантических ролей Места, Времени: когда и Агенса.

▪ Место

Среди наиболее типичных способов реализации семантической роли Места в корпусном материале можно выделить следующие несколько частично пересекающихся³¹ подгрупп:

- Указательные наречия места *там, тут, здесь* (19 вхождений)

(137) *и там был чудовищно замусоренный пляж* [КРУР]

- Именная группа с синтаксической структурой *В + Предложный падеж* (32 вхождения)

(138) *в советской школе этого было очень много* [КРУР]

³¹ Например, в материале корпуса есть вхождения типа *там свалка есть в Гдове* [КРУР].

- В том числе именная группа, содержащая имя собственное, которое относится к географическому объекту (15 вхождений)

(139) *Нечто похожее было и в Нижегородской области*
[НКРЯ]

- Именная группа с синтаксической структурой *На* + *Предложный падеж* (9 вхождений)

(140) *потому что он будет на лавиноопасном склоне.*
[НКРЯ]

▪ **Время: когда**

- Именная группа с синтаксической структурой *В* + *Предложный падеж* (16 вхождений)

(141) *первые наши поездки были в две тысячи одиннадцатом году* [КРУР]

- Наречия времени (*сейчас, раньше, недавно, потом* и т. д.) (13 вхождений)

(142) *сейчас в центре внимания будет молодёжная политика* [КРУР]

▪ **Агенс: способы выражения и проблема выделения**

Способы выражения семантической роли Агенса представляют особый интерес. Во всех случаях реализации семантической роли Агенса в корпусе соответствующие распространители имеют форму именной группы в именительном падеже, что, однако, не может являться критерием его выделения: такое же синтаксическое оформление имеют и большинство гораздо более частотных распространителей.

В отличие от распространителей, выполняющих роль *Время: когда* и *Места*, распространители со значением Агенса невозможно выявить автоматически на основе частотного словаря: тогда как, например, *здесь* и *раньше* могут быть интерпретированы алгоритмом, как выполняющие семантическую роль *Время: когда* и *Места* соответственно, местоимение и одушевленное существительное может с такой же (и, в рамках корпуса клауз с глаголом *быть*, даже с гораздо большей) вероятностью выражать значение Экзистенциального субъекта, Объекта Обладания или Тождественности.

На материале корпуса можно проследить только четыре тенденции в способах реализации Агенса. Три из них были описаны в п. 3.2. и относятся к склонности Агенса сочетаться с семантическими ролями Роли и Места, а также с неспособностью сочетаться с семантической ролью Тождественности.

В контексте же способов кодирования значения можно заметить, что во всех четырех вхождениях с распространителями, которые являются именами собственными в Именительном падеже, реферирующими к конкретным личностям, этот распространитель выполняет роль Агенса:

(143) *И **Явлинский** (Агенс) был коммунист (Роль)* [НКРЯ]

(144) *Когда недавно **Олег Медведев** (Агенс)*

был в Казани и там начал рассказывать [НКРЯ]

Кроме того, в достаточно частотных случаях сочетания Агенса с Ролью (формально совпадающей с Функцией, которая также типично выражается именно группой в творительном падеже без предлога), все случаи, где слот Роли/Функции занят названием профессии или неким другим обозначением групповой принадлежности (см. пример 142 выше), Слот Агенса/Экзистенциального субъекта оказывается заполнен именно Агенсом.

Итак, комплексный анализ синтаксической структуры клаузы вместе с данными о наиболее частотных случаях сочетаемости семантических ролей позволяет разрешить некоторые случаи неоднозначности, в том числе и автоматически.

Тем не менее, очевидно, что вопрос выделения критериев для автоматического выделения Агенса в клаузах с глаголом *быть* остается на данный момент нерешенным.

Заключение

В данной работе было представлено исследование, посвященное анализу клауз с формами глагола *быть* в русской устной речи и описанию его результатов в формате, который будет способствовать работе алгоритма автоматического распознавания русской устной речи.

В первой главе работы были проанализированы и обобщены существующие классификации семантических ролей с тем, чтобы на их основе создать опорный инвентарь семантических ролей, на который можно было бы ориентироваться при создании более специфических инвентарей, адаптированных под нужды конкретного исследования. Затем этот инвентарь был расширен и дополнен способами выражения семантических ролей, так чтобы более точно описывать именно русскую устную речь (результат представлен в Приложениях 1, 2).

В следующей части работы был проведен анализ клауз с глаголом *быть*, в результате чего был сформирован модифицированный инвентарь семантических ролей, с помощью которого можно описать специфику значений и функционирования в русской устной речи именно форм глагола *быть*. Так, в «репертуар» семантических ролей были введены следующие частотные в материале корпуса семантические роли:

- Экзистенциальный субъект
- Объект обладания
- Посессор
- Тождественность
- Роль
- Функция

Также был обсужден ряд неоднозначных случаев и предложены критерии для их разрешения.

Далее корпусный материал был описан в терминах перечисленных выше (а также и более традиционно выделяемых) ролей, в результате чего были

сделаны выводы о частотности активации семантических ролей глагола *быть* в русской устной речи, их распределении по формам глагола *быть*, тенденциях их сочетаемости в рамках одной клаузы и основных способах кодирования наиболее частотных из них.

Так, можно сделать следующие общие наблюдения о клаузах с глаголом *быть* в проанализированном материале русской устной речи:

- значения глагола *быть* обуславливают специфику инвентаря семантических ролей, выполняемых присоединяемыми им распространителями;
- так, полное описание клауз с формами глагола *быть* требует введения особых семантических ролей, малочастотных для распространителей среднего глагола русского языка и высокочастотных для распространителей глагола *быть*;
- самыми частотными семантическими ролями оказываются Экзистенциальный субъект, Место, Посессор и Объект Обладания;
- клаузы с глаголом *быть* демонстрируют большую вариативность в возможностях сочетания в их пределах семантических ролей, однако прослеживается также и ряд общих закономерностей:
 - ✓ наиболее типичной оказывается ситуация сочетаемости в пределах одной клаузы двух семантических ролей;
 - ✓ в материале корпуса встречаются пары семантических ролей, чрезвычайно склонные к сочетаемости друг с другом;
- в разной степени частотные семантические роли Экзистенциальный субъект, Объект обладания, Агенс и Тождество реализуются формально схожим образом (именные группы в именительном или родительном падеже) и, следовательно, представляют наибольшую сложность для алгоритма распознавания устной речи;
- в некоторых (но не во всех) случаях можно выявить закономерности сочетаемости этих ролей с другими, которые могут помочь при попытке интерпретировать семантическую роль распространителя.

Список использованной литературы

Научная литература

1. Апресян В. Ю. и др. Активный словарь русского языка. М.: Языки славянской культуры, 2014.
2. Апресян Ю. Д. Синонимические средства языка //Избранные труды. М.: Языки рус. культуры: Восточная литература. 1995.
3. Богуславский И. М. О противопоставлении актантов и сирконстантов в Московской семантической школе //Prace Filologiczne. – 2020. – №. 75/1. – С. 53–70-53–70.
4. Богуславский И. М. Сфера действия лексических единиц. М., 1996.
5. Гак В. Г. К проблеме синтаксической семантики // Инвариантные синтаксические значения и структура предложения / Ред. Н. Д. Арутюнова. М., 1969.
6. Годгильдиева М. М., Захаров В. П. Создание словаря валентности русского языка на основе компьютерного словаря валентности чешского языка Verbalex // Труды объединённой научной конференции "Интернет и современное общество". – 2016. – С. 14-23.
7. Евгеньева А. П. Малый академический словарь //М.: Институт русского языка Академии наук СССР. – 1957. – Т. 1984. – С. 698.
8. Жолковский А. К. О правилах семантического анализа //Машинный перевод и прикладная лингвистика: Сб. науч. трудов МГПИИЯ. – 1964. – №. 8. – С. 17-32.
9. Кацнельсон С. Д. О грамматической категории //Вестник ЛГУ. – 1948. – Т. 2. – С. 114-134.
10. Кацнельсон, С. Д. К понятию типов валентности / С. Д. Кацнельсон // Вопросы языкознания. — 1987. — № 3. — С. 20—32.
11. Кашкин Е. В., Ляшевская О. Н. Семантические роли и сеть конструкций в системе FrameBank //Компьютерная лингвистика и интеллектуальные

технологии. По материалам ежегодной международной конференции «Диалог. – 2013. – Т. 1. – С. 297-311.

12. Крылов С. А. Семантическая роль как элемент метаязыков общей и специальной типологии //40 лет Санкт-Петербургской типологической школе. – 2004. – С. 233-252.

13. Леонтьева Н. Н. Семантический анализ и смысловая неполнота текста: Автореф. дис.... канд. филол. наук. М., 1968.

14. Лингвистический энциклопедический словарь/ Главный редактор В. Н. Ярцева. М.: Советская энциклопедия, 1990. [Электронный ресурс]. URL: <http://tapemark.narod.ru/les/079c.html>

15. Мельчук И. А. Опыт теории лингвистических моделей «Смысл \Leftrightarrow Текст». М.: Наука, 1974.

16. Мельчук И. А., Холодович А. А. Залог. (Определение. Исчисление) //Народы Азии и Африки. – 1970. – №. 4. – С. 111-124.

17. Муравенко, Е. В. О случаях нетривиального соответствия семантических и синтаксических валентностей глагола // Семиотика и информатика 36. М.Ж Языки русской культуры, 1998.

18. Ляшевская, О. Н., Шаров, С. А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009.

19. Плунгян В. А. Общая морфология: Введение в проблематику. – 2012.

20. Плунгян В. А. Введение в грамматическую семантику: грамматические значения и грамматические системы языков мира. Москва: Российский государственный гуманитарный университет, 2011. – 672 с.

21. Риехакайнен Е. И. Непереходные глаголы движения в ментальном лексиконе носителя русского языка //Социо- и психолингвистические исследования, 2015. – №. 3.

22. Теньер Л. Основы структурного синтаксиса. / Пер. с франц. Вступ. ст. и общ. ред. В. Г. Гака. — М.: Прогресс, 1988. — 656 с.

23. Тестелец Я. Г. Введение в общий синтаксис. – Федеральное государственное бюджетное образовательное учреждение высшего образования Российский государственный гуманитарный университет, 2001.
24. Храковский В. С. Понятие сирконстанта и его статус //Семиотика и информатика. – 1998. – Т. 36. – С. 141-153.
25. Bisang W. et al. (ed.). Valency: Theoretical, descriptive and cognitive issues. – Mouton de Gruyter, 2007.
26. Chomsky N. Lectures on Government and Binding (Dordrecht: Foris) //Studies in generative grammar. – 1981. – Т. 9.
27. Comrie B. Language universals and linguistic typology: Syntax and morphology. – University of Chicago press, 1989.
28. de Groot, A. W. 1964 Reprint. Structurele Syntaxis. The Hague: Servire. Original edition, The Hague: Servire, 1949.
29. Dik S. C. Stepwise lexical decomposition. – 1978.
30. Emons R. Valency grammar in mind //Valency. – De Gruyter Mouton, 2008. – С. 183-192.
31. Fillmore C. J. Types of lexical information //Studies in syntax and semantics. – Springer, Dordrecht, 1969. – С. 109-137.
32. Fillmore Ch. J. Frame semantics, Linguistics in the morning calm: Selected papers from the SICOL-1981, Hanship, Seoul, 1982. Ppp. 111–137
33. Fillmore Ch. J. The case for case reopened, in Cole P., Sadock J. M. (eds.), Grammatical Relations, Acad. Press, New York, 1977. Pp. 59–81.
34. Fillmore Ch. J. The Case for Case, in Bach E. and Harms (Ed.), Universals in Linguistic Theory. New York, 1968. Pp. 1–88. 20.
35. Haspelmath M., Müller-Bardey T. Valency change //2. Halbband. – De Gruyter Mouton, 2008. – С. 1130-1145.
36. Hlavácková D., Horák A. Verbalex–new comprehensive lexicon of verb valencies for czech //Proceedings of the Slovko Conference. – 2005.

37. Hlaváčková D., Horák A., Kadlec V. Exploitation of the Verbalex verb valency lexicon in the syntactic analysis of Czech //International Conference on Text, Speech and Dialogue. – Springer, Berlin, Heidelberg, 2006. – C. 79-85.
38. Lyons J. A note on possessive, existential and locative sentences //Foundations of language. – 1967. – C. 390-396.
39. Lyons J. Towards a'notional'theory of the'parts of speech' //Journal of linguistics. – 1966. – T. 2. – №. 2. – C. 209-236.
40. Martin H., Müller-Bardey T. Valence change //Morphology: A handbook on inflection and word formation. –1991. – T. 2. – C. 1130-1145.
41. Matthews P. The scope of valency in grammar //Valency. – De Gruyter Mouton, 2007. – C. 3-14.
42. Mel'cuk. I. A. Actants in Semantics and Syntax I //Linguistics. – T. 42. – №. 1, 2004. – C. 1-66.
43. Mosel U. Towards a typology of valency //Arbeiten des Kölner Universalienprojekts, Institut für Sprachwissenschaft. – 1984. – T. 58. – C. 1-19.
44. Panevova J. 7. Valency Frames and the Meaning of the Sentence //The Prague School of Structural and Functional Linguistics. – 1994. – T. 41. – C. 223.
45. Panevová J. Inner participants and free adverbials //Prague Studies in Mathematical Linguistics. – 1978. – T. 6. – C. 227-254.
46. Somers H. L. On the validity of the complement-adjunct distinction in valency grammar. – 1984.
47. Thompson C. A., Levy R., Manning C. D. A generative model for semantic role labeling //European Conference on Machine Learning. – Springer, Berlin, Heidelberg, 2003. – C. 397-408.
48. Van Valin Jr R. D. et al. An introduction to syntax. – Cambridge University Press, 2001.
49. Van Valin, R.D. Jr. (ed). 1993. Advances in role and reference grammar. Amsterdam: Benjamins.

Электронные источники

1. Корпус русской устной речи (КРУР). URL: <http://russpeech.spbu.ru/> (дата обращения: 25.05.2021).
2. Национальный корпус русского языка (НКРЯ) URL: <http://ruscorpora.ru/> (дата обращения: 25.05.2021).
3. Русская грамматика / Н. Ю. Шведова (гл. ред.). — М.: Наука, 1980. URL: <http://rusgram.narod.ru/> (дата обращения: 25.05.2021).

Приложение 1. Опорный инвентарь семантических ролей и способов их выражения

Предлог	Падеж	пример	Семантическая роль
Без	Р. П.	<i>нельзя пройти без билета</i>	Причина
Без предлога	Им. П.		Агенс / субъект
Без предлога	Вин. П.	<i>Они приветствовали нас</i>	Адресат
Без предлога	Дат. П.	<i>рассказали нам</i>	Адресат
Без предлога	Дат. П.	<i>Девочка отдает собаку брату</i>	Бенефактив
Без предлога	Тв. П.	<i>шить иголкой</i>	Инструмент
Без предлога	Вин. П.	<i>Девочка отдает собаку брату</i>	Пациент / объект
Без предлога	Им. П.	<i>дождь затопил посевы, страх гнал его в путь</i>	Причина
Без предлога	Дат. П.	<i>царь обрадовался известию</i>	Причина
Без предлога	Тв. П.	<i>гордиться котом</i>	Причина
Без предлога	Тв. П.	<i>стать котом</i>	Роль/ функция
Без предлога	Тв. П.	<i>красить охрой</i>	Средство
Без предлога	Им. П.	<i>Читателям понравился роман</i>	Стимул
Без предлога	Вин. П.	<i>Мальчик видит котика</i>	Стимул
Без предлога	Тв. П.	<i>увлекаться, интересоваться котами</i>	Стимул
Без предлога	Дат. П.	<i>Читателям понравился роман</i>	Экспериенцер
Без предлога	Дат. П.	<i>сопротивляться кому</i>	Контрагент
Благодаря	Дат. П.	<i>благодаря котикам</i>	Причина
В	Предл. П.	<i>в доме</i>	Место
В	Предл. П.	<i>прочитать в книге (в знач. Узнать)</i>	Источник
В	Предл. П.	<i>надежный в работе</i>	Аспект/ Сфера
В	Предл. П.	<i>в августе, в детстве</i>	Время: когда
В	Вин. П.	<i>идти в школу</i>	Направление
В	Предл. П.	<i>девушка в шляпе, ходить в очках</i>	Определение
В	Предл. П.	<i>кричать в гневе</i>	Состояние
Вдоль	Р. П.	<i>вдоль дороги</i>	Место
Вдоль	Р. П.	<i>вдоль дороги</i>	Маршрут

Вместо	Р. П.	<i>вместо тетради</i>	Замещение
Внутри	Р. П.		Место
Вокруг	Р. П.	<i>Обойти вокруг сада</i>	Маршрут
Вокруг	Р. П.	<i>раскиданы вокруг дома</i>	Место
Вопреки	Дат. П.		Контрагент
Для	Р. П.	<i>купить для друга</i>	Бенефактив
Для	Р. П.	<i>купить для похода</i>	Цель
До	Р. П.	<i>дойти до дома</i>	Направление
До	Р. П.	<i>до вечера</i>	Время: конечная точка
За	Тв. П.	<i>за столом</i>	Место
За	Тв. П.	<i>за обедом</i>	Время: когда
За	Вин. П.	<i>Сесть за стол</i>	Направление
Из	Р. П.		Источник
Из	Р. П.	<i>сделать из принципа</i>	Причина
Из-за	Р. П.	<i>выйти из-за леса</i>	Источник
Из-за	Р. П.	<i>пропустить из-за болезни</i>	Причина
Из-под	Р. П.	<i>выплыть из-под моста</i>	Источник
К	Дат. П.	<i>прийти к часу дня, закончить к вечеру</i>	Время: конечная точка
К	Дат. П.	<i>плыть к берегу</i>	Направление
К	Дат. П.	<i>прикрепить к доске, подойти к дому</i>	Направление
Кроме	Р. П.	<i>кроме моих друзей, никто не знал об этом</i>	Исключение
Между	Тв. П.	<i>между садом и школой</i>	Место
Между	Тв. П.	<i>между занятиями</i>	Время: когда
Между	Тв. П.	<i>сходство между предметами</i>	Объекты сравнения
Мимо	Р. П.	<i>пройти мимо музея</i>	Маршрут
Мимо	Р. П.	<i>попасть мимо мишени</i>	Направление
На	Предл. П.	<i>находится на крыше</i>	Место
На	Предл. П.	<i>на этой неделе, на днях, на рассвете</i>	Время: когда
На	Вин. П.	<i>положить на стол</i>	Направление
На	Предл. П.	<i>ехать на велосипеде</i>	Определение: как / способ
На	Предл. П.	<i>ходить на руках</i>	Определение: как / способ
На	Предл. П.	<i>строить на фундаменте, жарить</i>	Определение: как / способ

		<i>на масле, говорить на языке, блины на молоке</i>	
На	Предл. П.	<i>на ходу, на лету</i>	Определение: как / способ
Навстречу	Дат. П.	<i>идти навстречу опасностям</i>	Направление
Навстречу	Дат. П.	<i>машины двигались навстречу друг другу</i>	Направление
Над	Тв. П.	<i>над столом</i>	Место
Накануне	Р. П.		Время: когда
Наперекор	Дат. П.	<i>наперекор судьбе</i>	Контрагент
Напротив	Р. П.	<i>напротив дома</i>	Место
Напротив	Р. П.	<i>поставить напротив стола</i>	Точка отсчета
Недалеко от	Р. П.		Место
О	Вин. П.	<i>споткнуться о камень, разбиться о скалы</i>	Инструмент
О	Предл. П.		Пациент / объект
О	Предл. П.		Содержание
О	Предл. П.		Содержание
Около	Р. П.		Место
От	Р. П.		Источник
От	Р. П.	<i>дрожать от холода</i>	Причина
Перед	Тв. П.		Место
Перед	Тв. П.	<i>перед занятиями</i>	Время: когда
По	Вин. П.	<i>сыт по горло, нога по колено</i>	Предел
По	Дат. П.	<i>по выходным</i>	Время: когда
По	Дат. П.	<i>плыть по воде</i>	Место
По	Дат. П.	<i>по разным городам, по миру</i>	Место
По	Дат. П.	<i>инженер по технике безопасности, учебник по русскому</i>	Аспект/ Сфера
По	Дат. П.	<i>ошибиться по неопытности</i>	Причина
По	Вин. П.	<i>по декабрь</i>	Время: конечная точка
По	Дат. П.	<i>послать по почте</i>	Определение: как / способ*
По	Предл. П.	<i>по приезде</i>	Время: когда
Под	Тв. П.	<i>гулять под дождем, находиться по снегом</i>	Место
Под	Вин. П.	<i>уйти под воду</i>	Направление
Под	Вин. П.	<i>под барабанный бой</i>	Условия

После	Р. П.	<i>после похода</i>	Время: когда
При	Предл. П.	<i>жить при Иване Грозным, при встрече</i>	Время: когда
При	Предл. П.	<i>добиться при желании, при луне</i>	Условия
При	Предл. П.	<i>сказать при постороннем, сделать при свидетелях</i>	Условия
При	Предл. П.	<i>при любой погоде; при всем том</i>	Условия
Против	Р. П.		Контрагент
С	Р. П.	<i>работать с утра</i>	Время: начальная точка
С	Тв. П.	<i>проснуться с рассветом</i>	Время: когда
С	Р. П.	<i>взять с полки</i>	Источник
С	Тв. П.	<i>комната с высоким потолком, корзина с грибами</i>	Определение
С	Р. П.	<i>сказать со злости</i>	Причина
С	Тв. П.	<i>разговаривать с преподавателем, приехать с портфелем</i>	Совместность
С	Тв. П.	<i>встретиться с другом, знакомиться с..</i>	Совместность
С	Тв. П.	<i>слушать с интересом, ждать с нетерпением</i>	Определение: как / способ*
С	Тв. П.	<i>рыбачить с удочкой, подойти с улыбкой</i>	Определение: как / способ*
С	Тв. П.	<i>как с этим быть</i>	Определение: как / способ*
С	Тв. П.		Определение: как / способ*
Сквозь	Вин. П.	<i>видеть сквозь стекло</i>	Определение: как / способ*
Снаружи	Р. П.		Место
Согласно	Дат. П.	<i>согласно сведениям...</i>	Источник
Согласно	Дат. П.	<i>согласно распоряжению..</i>	Причина
Среди	Р. П.	<i>жить среди людей</i>	Место
Среди	Р. П.	<i>придти среди ночи</i>	Время: когда
У	Р. П.	<i>ждать у входа, жить у тети, работать у соседей</i>	Место
У	Р. П.	<i>узнать у одноклассника, взять у соседа</i>	Источник
У	Р. П.		Посессор

Через	Вин. П.	<i>мост через...</i>	Маршрут
--------------	---------	----------------------	----------------

Приложение 2. Опорный инвентарь семантических ролей и способов их выражения: влияние конструкций

Конструкция	Предлог	Падеж	пример	Семантическая роль	Структура
Хорошо, плохо, больно и др. обозначения состояния	Без предлога	Дат. П.	<i>Мне плохо</i>	Экспериенцер	Сущ./местоимение в Дат. П. + состояние
Можно, нельзя	Без предлога	Дат. П.	<i>мне можно делать...</i>	Экспериенцер	
Слева, справа, далеко, недалеко, близко	От	Р. П.	<i>слева от дома</i>	Точка отсчета	слева от, справа от + сущ в Р.п.
Рядом	С	Тв. П.	<i>рядом с домом</i>	Место	
Числительные (2, 3, 4)		Р. П.		Агенс / субъект	2, 3, 4 + сущ в Р.п. ед. ч.
Числительные (5, 10, 20...)		Тв. П.	<i>город был построен котом</i>	Агенс / субъект	5, 10, 20 + сущ в Р.п. мн. ч.
Сколько, много, немного, мало, несколько		Им. П.		Объект обладания	сколько, много, немного + сущ в Р.п. мн. ч.
Прилагательные и наречия в сравнительной степени		Р. П.	<i>2 лавины разрушили деревню</i>	Причина	Прил. Нар. В сравнительной степени + сущ в Р.п. мн. ч.
Числительные (2, 3, 4)		Р. П.		Причина	2, 3, 4 + сущ в Р.п. ед. ч.
Числительные (5, 10, 20...)		Тв. П.	<i>лавинной были уничтожены посевы</i>	Причина	5, 10, 20 + сущ в Р.п. мн. ч.
Сколько, много, немного, мало, несколько		Р. П.	<i>он увидел 3 человека</i>	Стимул	сколько, много, немного... + сущ в Р.п. мн. ч.
Числительные (2, 3, 4)		Р. П.		Стимул	2, 3, 4 + сущ в Р.п. ед. ч.
Числительные (5, 10, 20...)		Тв. П.	<i>люди были впечатлены котом</i>	Стимул	5, 10, 20 + сущ в Р.п. мн. ч.

Сколько, много, немного, мало, несколько	За	Вин. П.	<i>за десять рублей</i>	Цена / Стоимость / Обмен	сколько, много, немного... + сущ в Р.п. мн. ч.
--	-----------	------------	-----------------------------	---	--

Приложение 3. Распределение по формам глагола *быть*: Экзистенциальный субъект

		Форма	Количество вхождений
Инфинитив		быть	19
есть			92
Формы будущего времени	33	буду	0
		будем	0
		будет	29
		будешь	0
		будете	0
		будут	4
Формы прошедшего времени	99	был	18
		была	25
		было	29
		были	27
Формы императива	0	будь	1
		будьте	0
Итого		244	

Приложение 4. Распределение по формам глагола *быть*: Место

		Форма	Количество вхождений
Инфинитив		быть	4
есть			24
Формы будущего времени	8	буду	0
		будем	0
		будет	8
		будешь	0
		будете	0
		будут	0
Формы прошедшего времени	31	был	6
		была	4
		было	8
		были	13
Формы императива	0	будь	0
		будьте	0
Итого		67	

Приложение 5. Распределение по формам глагола *быть*: Посессор

		Форма	Количество вхождений
Инфинитив		быть	2
есть			26
Формы будущего времени	11	буду	0
		будем	0
		будет	11
		будешь	0
		будете	0
		будут	0
Формы прошедшего времени	23	был	5
		была	1
		было	12
		были	5
Формы императива	0	будь	0
		будьте	0
Итого		62	

Приложение 6. Распределение по формам глагола *быть*: Объект Обладания

		Форма	Количество вхождений
Инфинитив		быть	2
есть			25
Формы будущего времени	9	буду	0
		будем	0
		будет	9
		будешь	0
		будете	0
		будут	0
Формы прошедшего времени	25	был	5
		была	1
		было	14
		были	5
Формы императива	0	будь	0
		будьте	0
Итого		61	

Приложение 7. Распределение по формам глагола *быть*: Определение

		Форма	Количество вхождений
Инфинитив		быть	14
есть			5
Формы будущего времени	5	буду	0
		будем	0
		будет	4
		будешь	0
		будете	0
		будут	1
Формы прошедшего времени	25	был	6
		была	9
		было	2
		были	8
Формы императива	0	будь	0
		будьте	0
Итого		49	

**Приложение 8. Распределение: когда по формам глагола *быть*:
Время: когда**

		Форма	Количество вхождений
Инфинитив		быть	2
есть			3
Формы будущего времени	7	буду	0
		будем	0
		будет	7
		будешь	0
		будете	0
		будут	0
Формы прошедшего времени	30	был	6
		была	4
		было	11
		были	9
Формы императива	0	будь	0
		будьте	0
Итого		42	

Приложение 9. Распределение по формам глагола *быть*: Тождественность

		Форма	Количество вхождений
Инфинитив		быть	1
есть			10
Формы будущего времени	2	буду	0
		будем	0
		будет	2
		будешь	0
		будете	0
		будут	0
Формы прошедшего времени	15	был	4
		была	8
		было	1
		были	2
Формы императива	0	будь	1
		будьте	0
Итого		29	

Приложение 10. Распределение по формам глагола *быть*: Агенс

		Форма	Количество вхождений
Инфинитив		быть	6
есть			0
Формы будущего времени	4	буду	1
		будем	0
		будет	2
		будешь	0
		будете	0
		будут	1
Формы прошедшего времени	15	был	6
		была	3
		было	1
		были	5
Формы императива	0	будь	0
		будьте	0
Всего		25	